

O que é Aprendizado de Máquina (Machine Learning)?

Machine Learning

- Antes de entrarmos em Redes Neurais, Tensorflow, Keras API, etc... é uma boa ideia entender algumas ideias fundamentais sobre aprendizado de máquina.

Machine Learning

- Visão geral:
 - O que é Machine Learning
 - O que é Deep Learning?
 - Qual a diferença entre Aprendizado Supervisionado e Não Supervisionado
 - Processo de Aprendizagem Supervisionado
 - Avaliando o desempenho
 - Overfitting

O que é Machine Learning?

- O aprendizado de máquina é um método de análise de dados que automatiza a construção de modelos analíticos.
- Usando algoritmos que aprendem de forma iterativa com os dados, o aprendizado de máquina permite que os computadores encontrem informações ocultas, sem serem explicitamente programados onde procurar.

Para que é utilizado?

- Detecção de fraude.
- Resultados de pesquisas na web.
- Anúncios em tempo real em páginas da web.
- Pontuação de crédito.
- Previsão de falhas de equipamentos.
- Novos modelos de pricing.
- Detecção de intrusão de rede.
- Mecanismos de recomendação.
- Segmentação de clientes.
- Análise de sentimento de texto.
- A rotatividade de clientes.
- Reconhecimento de padrões e imagens.
- Filtragem de spam de e-mail.

O que são Redes Neurais (Neural Networks)?

- As Redes Neurais são uma maneira de modelar matematicamente sistemas de neurônios biológicos.
- Essas redes podem então ser utilizadas para resolver tarefas que muitos outros tipos de algoritmos não podem (por exemplo, classificação de imagens).
- Deep Learning simplesmente se refere a redes neurais com mais de uma camada oculta.

Machine Learning

- Existem diferentes tipos de aprendizado de máquina nos quais nos concentraremos:
 - Aprendizado Supervisionado
 - Aprendizado Não Supervisionado

Machine Learning

- Machine Learning
 - Modelos analíticos automatizados.
- Neural Networks
 - Um tipo de arquitetura de aprendizado de máquina modelada a partir de neurônios biológicos.
- Deep Learning
 - Uma rede neural com mais de uma camada oculta.

Machine Learning

- Vamos começar aprendendo sobre uma das tarefas mais comuns de aprendizado de máquina: **Aprendizado Supervisionado!**

Aprendizado Supervisionado (Supervised Learning)

Supervised Learning

- Os algoritmos de **aprendizado supervisionado** (supervised learning) são treinados usando exemplos **rotulados** (labeled), como uma entrada em que a saída desejada é conhecida.
- Por exemplo, um segmento de texto pode ter um rótulo de categoria, como:
 - **Spam** versus E-mail **Legítimo**
 - Análise **Positiva** versus **Negativa** do Filme

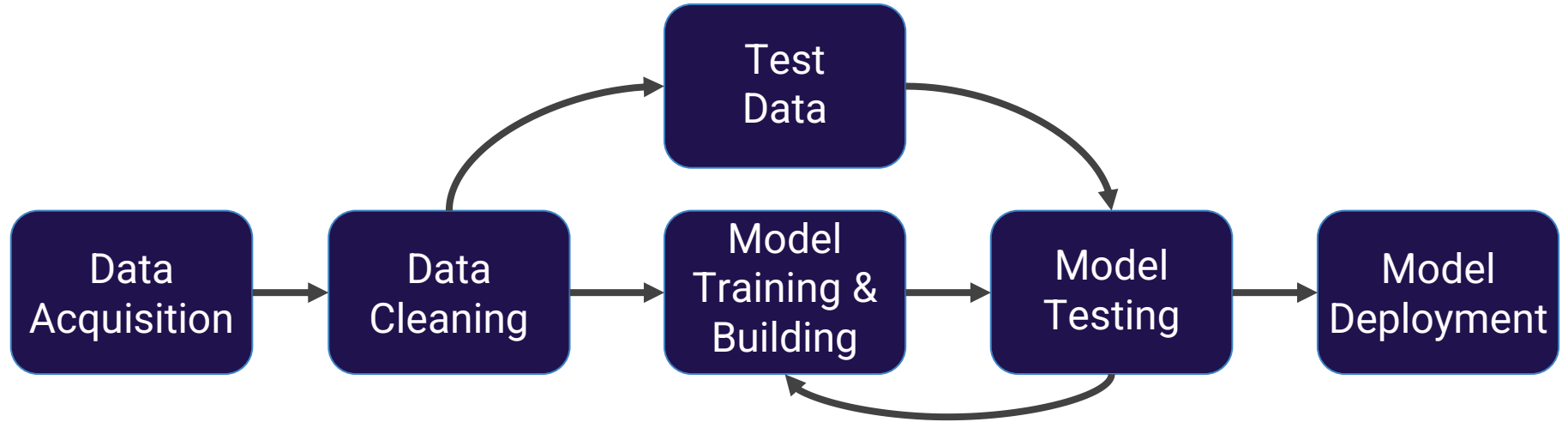
Supervised Learning

- A rede recebe um conjunto de entradas junto com as saídas corretas correspondentes, e o algoritmo aprende comparando sua saída real com as saídas corretas para encontrar erros.
- Em seguida, ele modifica o modelo de acordo.

Supervised Learning

- O aprendizado supervisionado é comumente usado em aplicativos em que dados históricos preveem eventos futuros prováveis.

Machine Learning Process



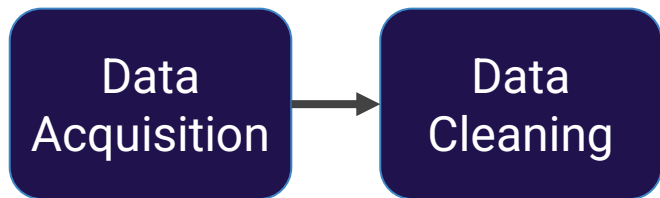
Machine Learning Process

- Obtenha seus dados! Clientes, Sensores, etc...

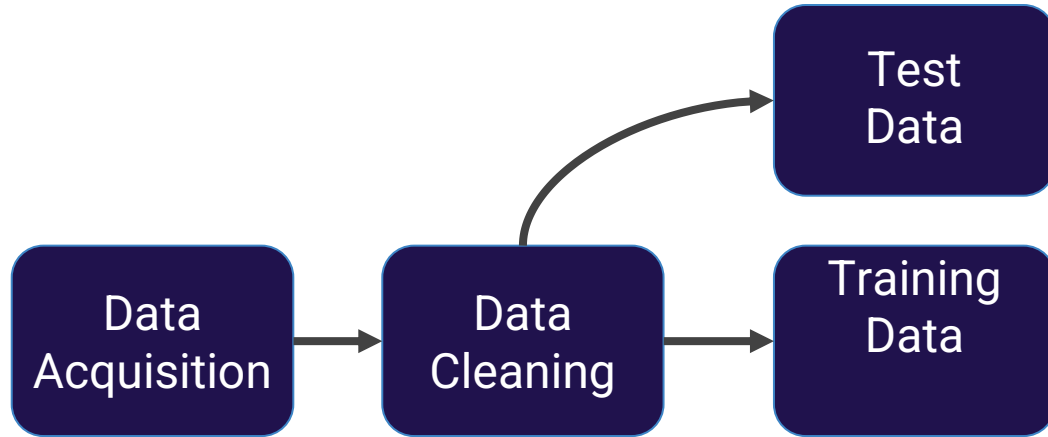
Data
Acquisition

Machine Learning Process

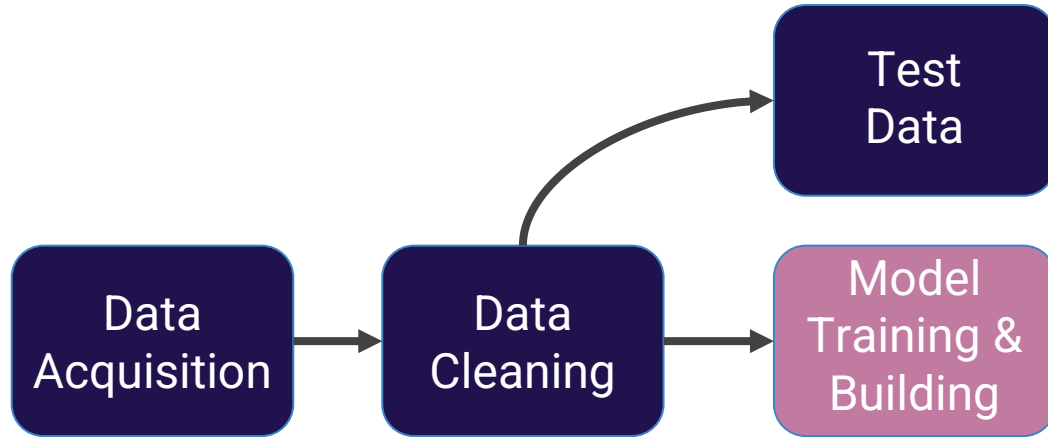
- Limpe e formate seus dados (utilizando Pandas)



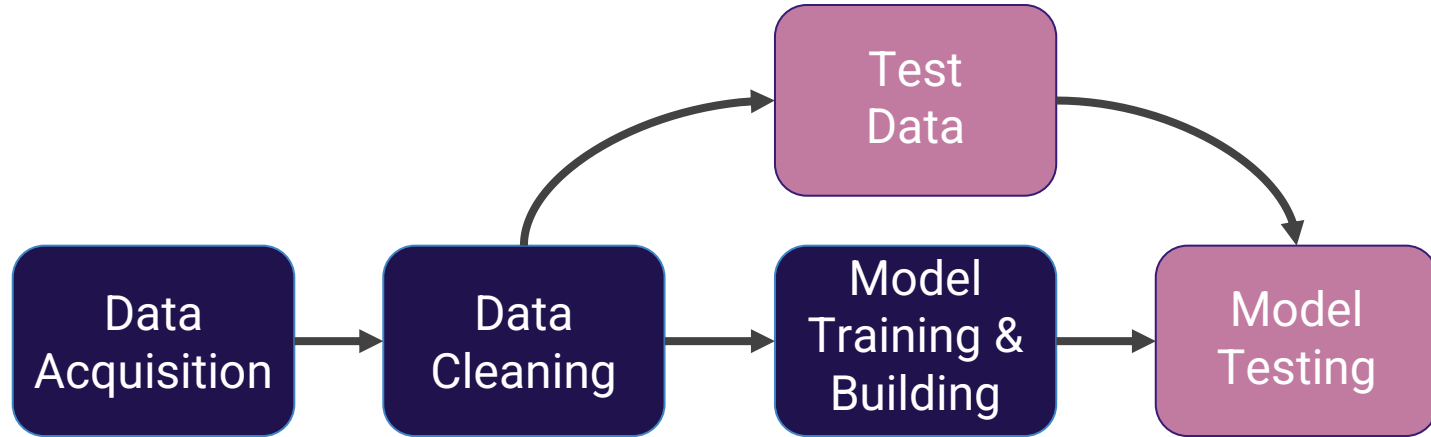
Machine Learning Process



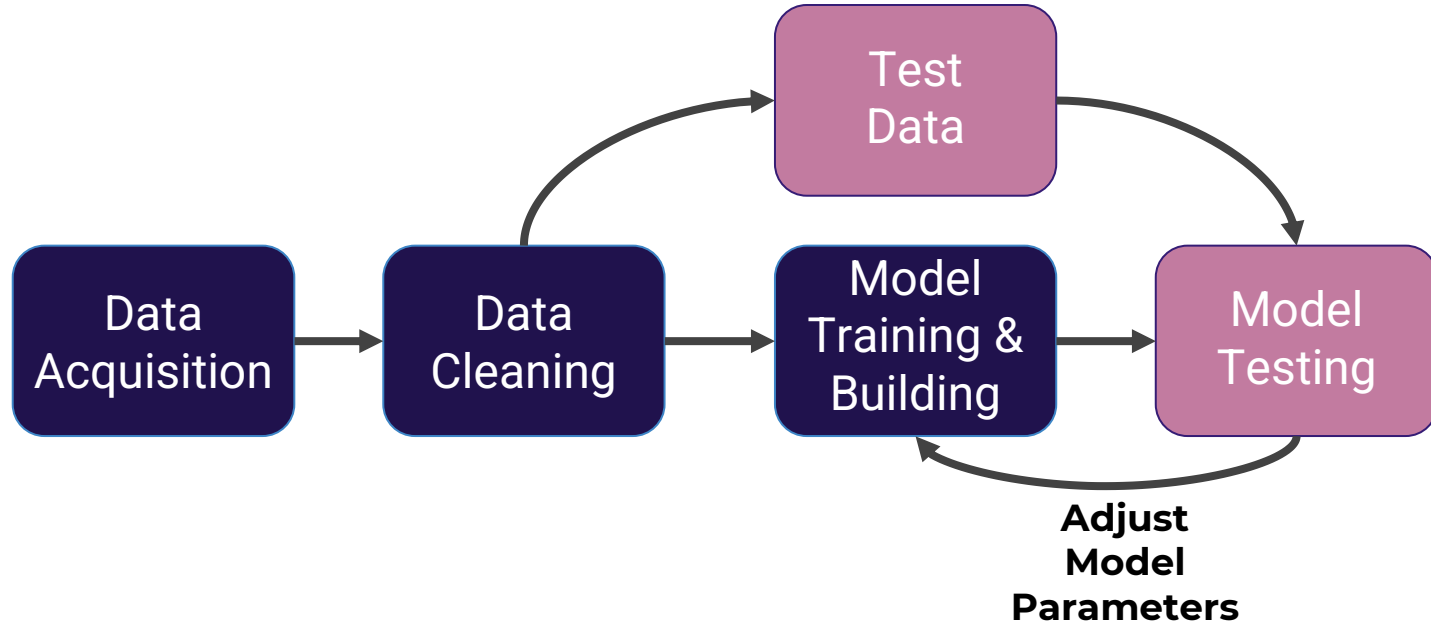
Machine Learning Process



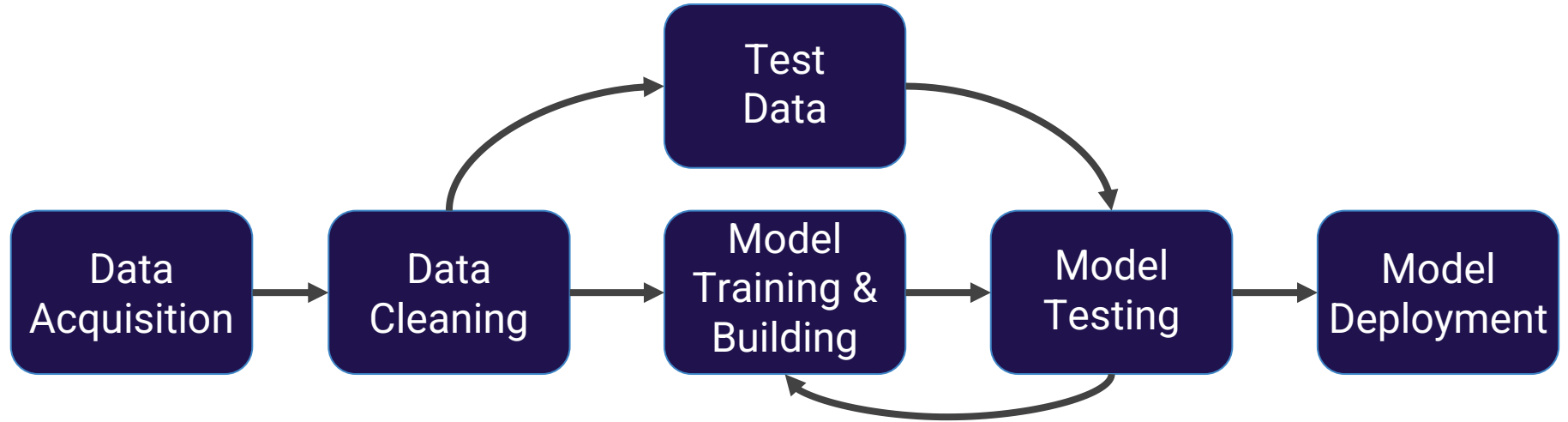
Machine Learning Process



Machine Learning Process



Machine Learning Process



Supervised Learning

- O que acabamos de ver é uma abordagem simplificada para o aprendizado supervisionado. Ela contém um problema!
- É justo usar nossa única divisão dos dados para avaliar o desempenho de nossos modelos?
- Afinal, tivemos a chance de atualizar os parâmetros do modelo repetidamente.

Supervised Learning

- Para corrigir esse problema, os dados geralmente são divididos em **3 conjuntos**
 - Training Data
 - Utilizado para treinar parâmetros do modelo
 - Validation Data
 - Utilizado para determinar quais hiperparâmetros do modelo devem ser ajustados
 - Test Data
 - Utilizado para obter alguma métrica de desempenho final

Supervised Learning

- Isso significa que depois de vermos os resultados no **conjunto de teste final** (final test set), não podemos voltar e ajustar nenhum parâmetro do modelo!
- Essa medida final é o que rotulamos como o verdadeiro desempenho do modelo.

Supervised Learning

- Simplificaremos nossos dados utilizando uma **divisão simples de treinamento/teste** (train/test split).
- Vamos simplesmente treinar e depois avaliar em um conjunto de teste (deixando a opção para os alunos voltarem e ajustarem os parâmetros).
- No futuro, você poderá facilmente realizar outra divisão para obter **3 conjuntos de dados**, caso desejar.

Sobreajuste e Subajuste (Overfitting and Underfitting)

Machine Learning

- Agora que entendemos o processo completo de aprendizado supervisionado, vamos abordar os tópicos importantes de **overfitting** e **underfitting**.

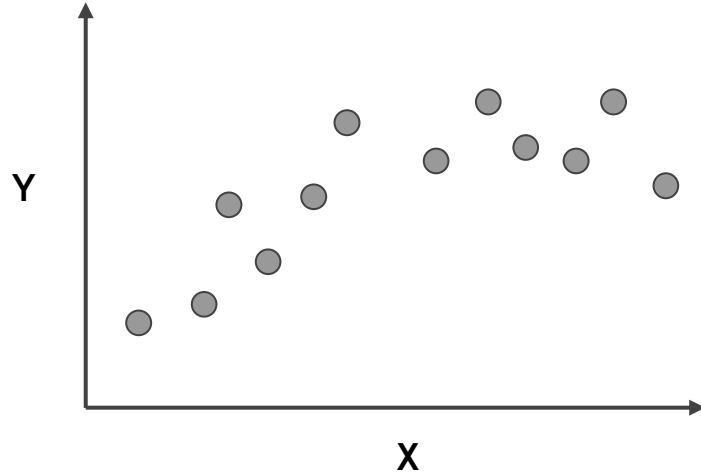
Machine Learning

- **Overfitting**

- O modelo se ajusta demais ao ruído dos dados.
- Isso geralmente resulta em **baixo erro nos conjuntos de treinamento** (training sets), **mas alto erro nos conjuntos de teste/validação** (test/validation sets).

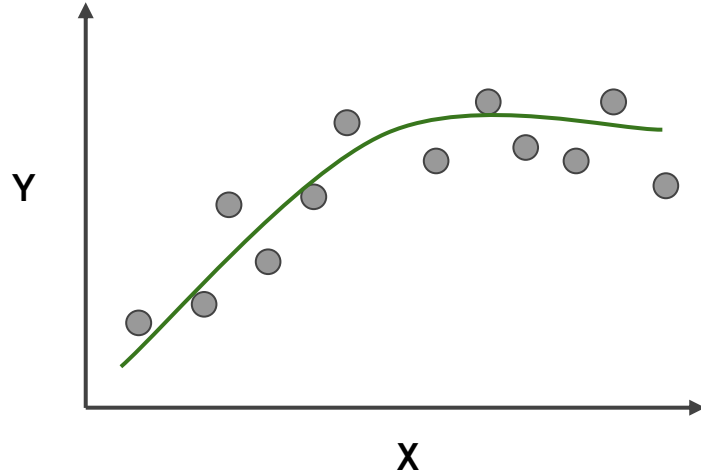
Machine Learning

Data



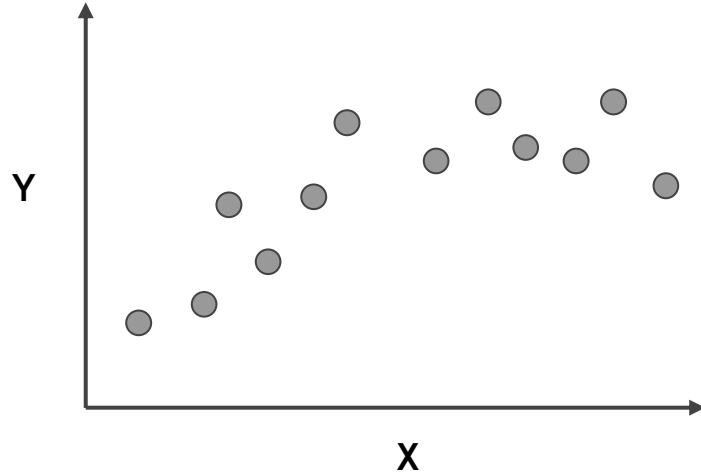
Machine Learning

Good Model



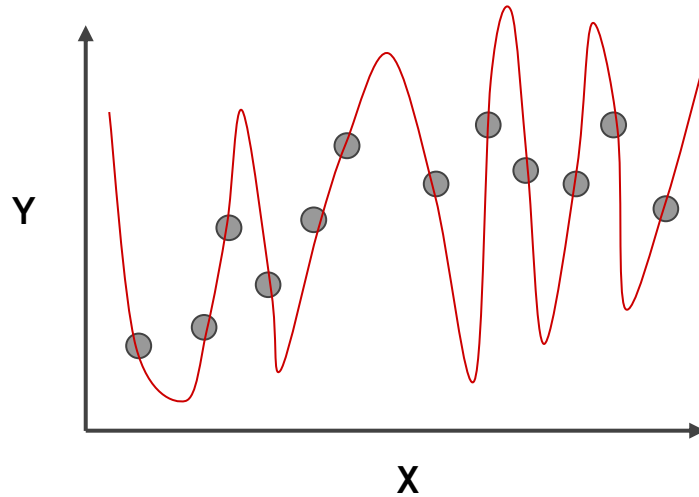
Machine Learning

- **Overfitting**



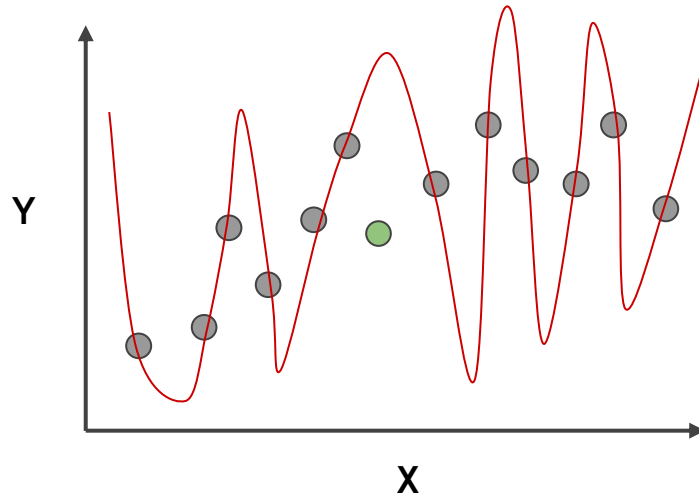
Machine Learning

- **Overfitting**



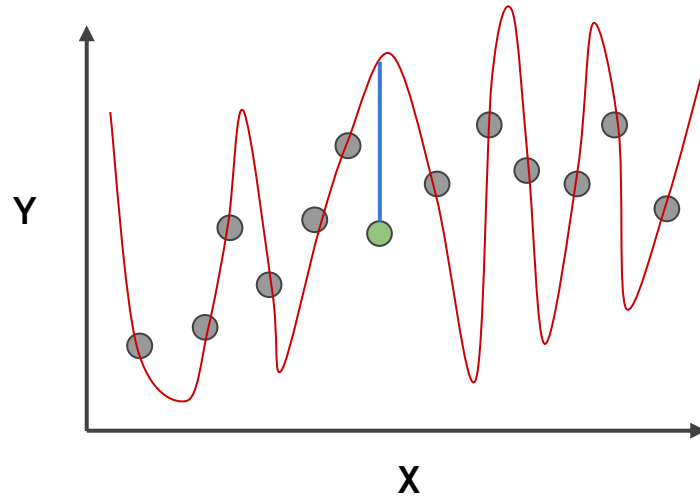
Machine Learning

- **Overfitting**



Machine Learning

- **Overfitting**



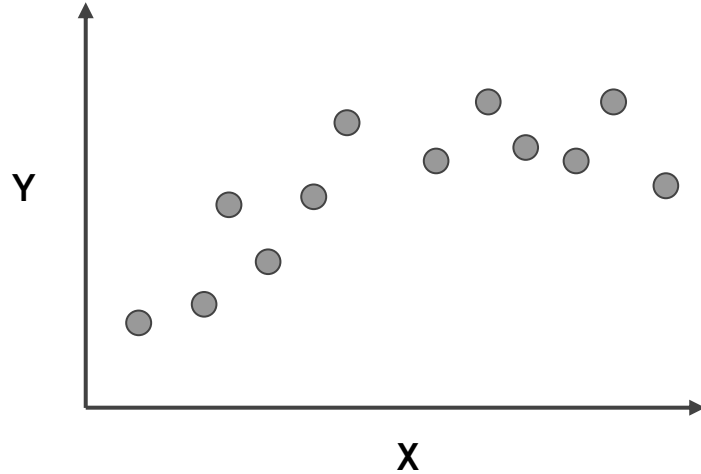
Machine Learning

- **Underfitting**

- O modelo não captura a tendência implícita dos dados e não se ajusta bem aos dados.
- Baixa variância, mas alto viés (bias).
- O underfitting é muitas vezes o resultado de um modelo excessivamente simples.

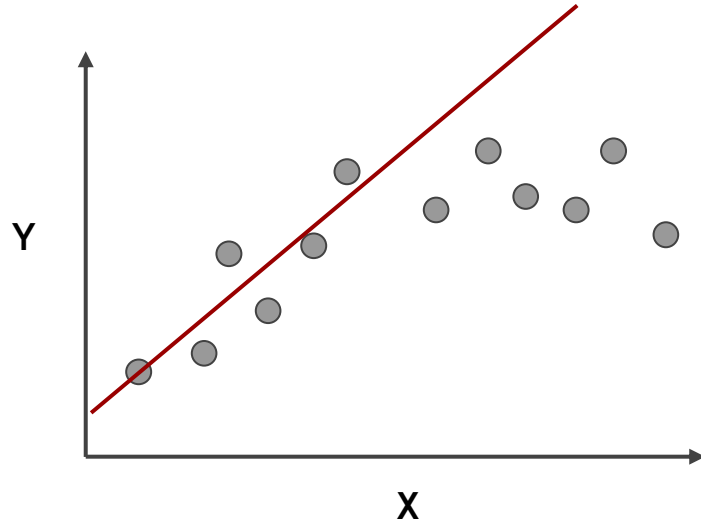
Machine Learning

Data



Machine Learning

Underfitting

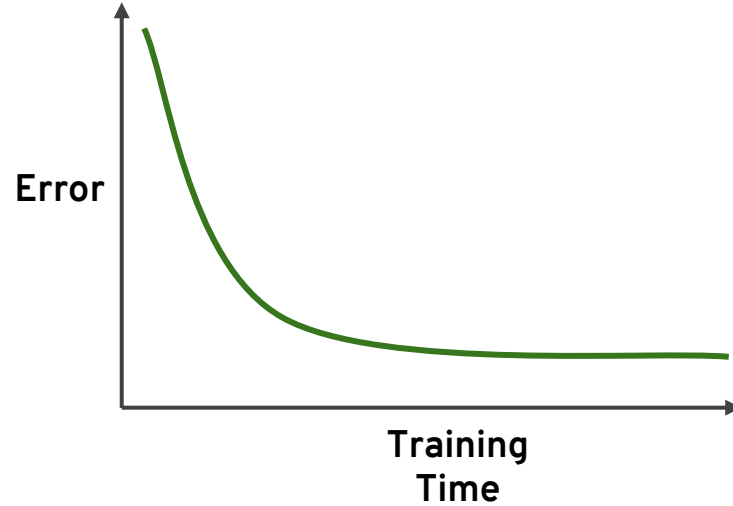


Machine Learning

- Esse conjunto de dados era fácil de visualizar, mas como podemos averiguar a existência de underfitting ou overfitting ao lidar com conjuntos de dados multidimensionais?
- Primeiro vamos imaginar que treinamos um modelo e depois medimos seu erro ao longo do tempo de treinamento.

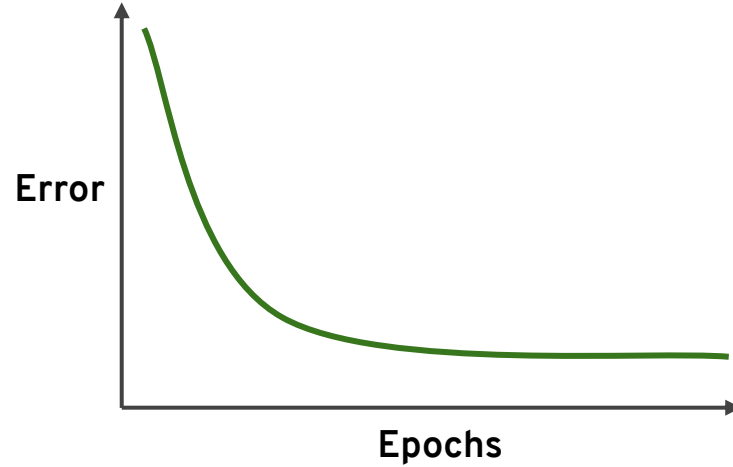
Machine Learning

- Good Model



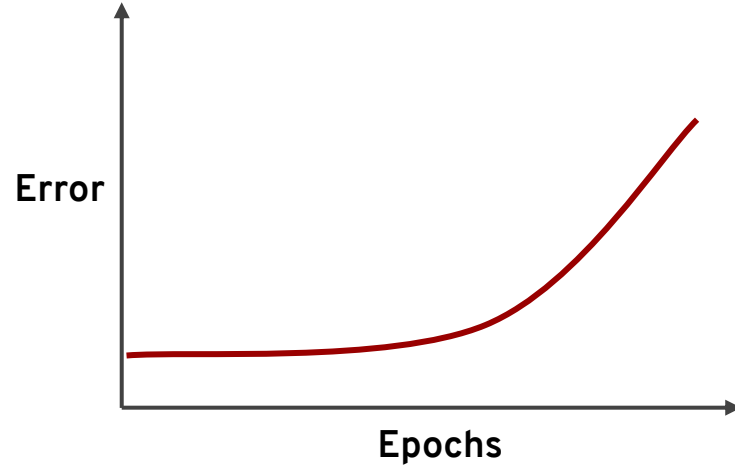
Machine Learning

- Good Model



Machine Learning

- Bad Model



Machine Learning

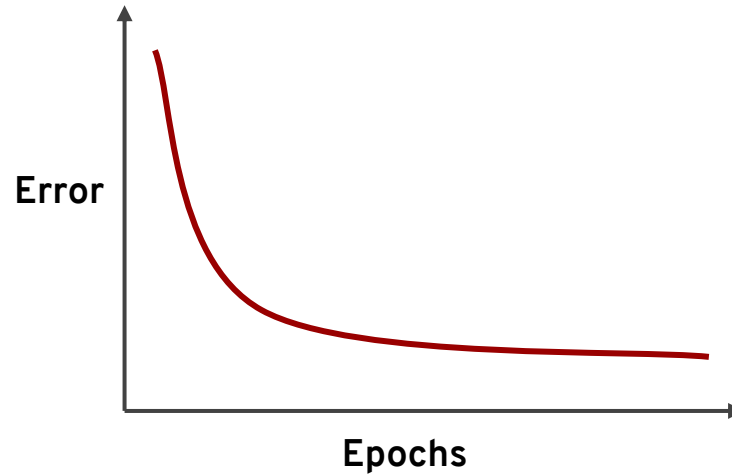
- Ao pensar em **overfitting** e **underfitting**, queremos verificar a relação do desempenho do modelo no conjunto de treinamento (training set) versus o conjunto de teste/validação (test/validation set).

Machine Learning

- Vamos imaginar que dividimos nossos dados em um **training set** e um **test set**

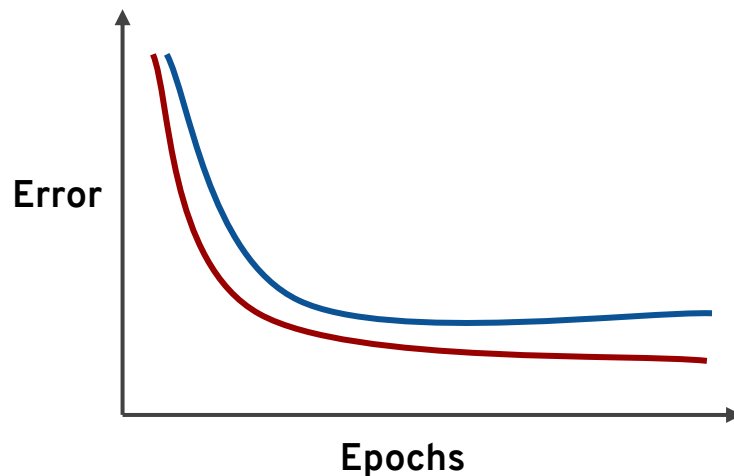
Machine Learning

- Primeiro vemos o desempenho no **training set**



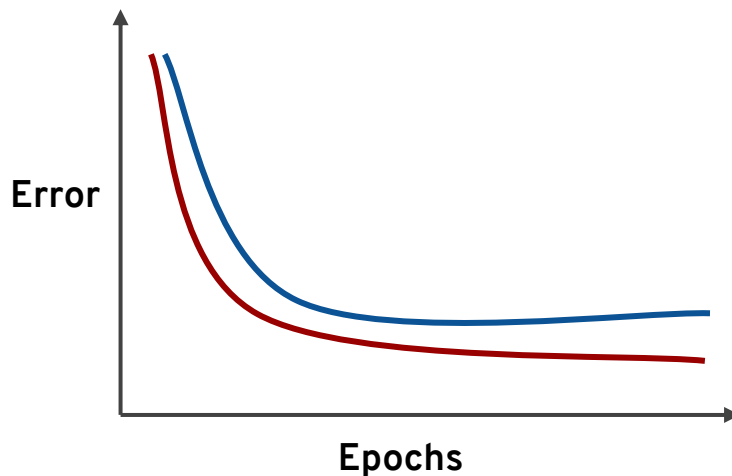
Machine Learning

- Em seguida, verificamos o desempenho no **test set**



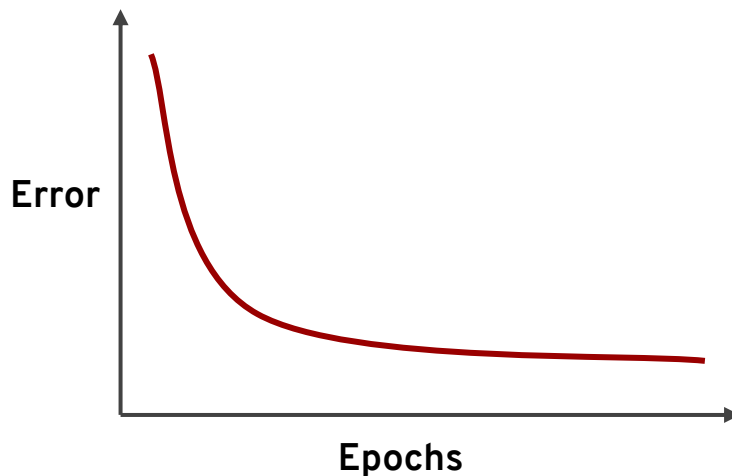
Machine Learning

- Idealmente, o modelo teria um bom desempenho em ambos, com comportamento semelhante.



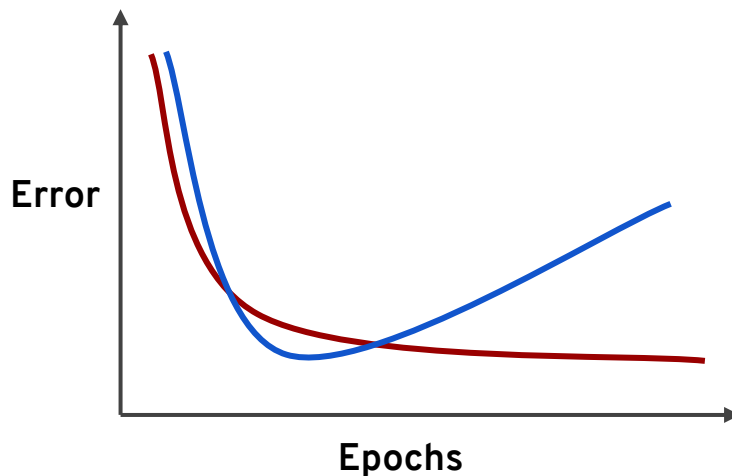
Machine Learning

- Mas o que acontece se ajustarmos demais (overfit) aos dados de treinamento? Isso significa que teríamos um desempenho ruim em novos dados de teste!



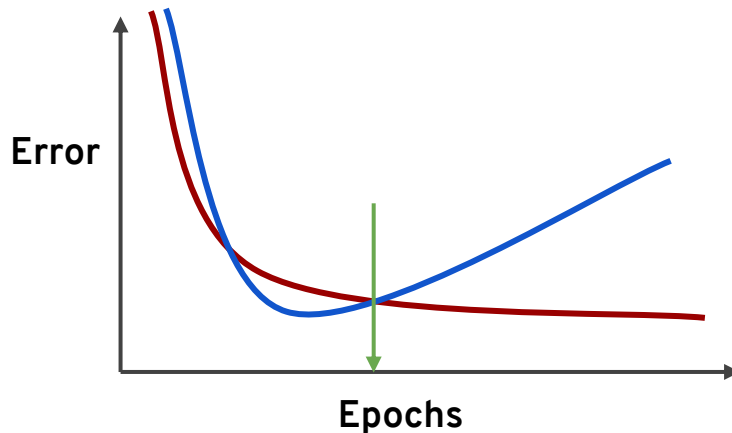
Machine Learning

- Mas o que acontece se ajustarmos demais (overfit) aos dados de treinamento? Isso significa que teríamos um desempenho ruim em novos dados de teste!



Machine Learning

- Esta é uma boa indicação de ter treinado demais nos dados de treinamento. Deve-se procurar um ponto para cortar o tempo de treinamento!



Machine Learning

- Vamos verificar essa ideia novamente quando começarmos a criar modelos!
- Por enquanto apenas fique atento a esse possível problema!

Avaliando o Desempenho (Evaluating Performance)

CLASSIFICAÇÃO
(CLASSIFICATION)

Model Evaluation

- Acabamos de saber que, após a conclusão do nosso processo de machine learning, usaremos **métricas de desempenho** (performance metrics) para avaliar o desempenho do nosso modelo.
- Vamos discutir as **métricas de classificação** (classification metrics) com mais detalhes!

Model Evaluation

- As principais **métricas de classificação** (classification metrics) que precisamos entender são:
 - Acurácia (Accuracy)
 - Revocação (Recall)
 - Precisão (Precision)
 - Pontuação-F1 (F1-Score)

Model Evaluation

- Mas primeiro, devemos entender o raciocínio por trás dessas métricas e como elas realmente funcionarão no mundo real!

Model Evaluation

- Normalmente, em qualquer tarefa de classificação, seu modelo pode alcançar apenas dois resultados:
 - Ou seu modelo estava **correto** em sua previsão.
 - Ou seu modelo estava **incorreto** na previsão.

Model Evaluation

- Felizmente, incorreto versus correto se expande para situações em que você tem várias classes.
- Para fins de explicação das métricas, vamos imaginar uma situação de **classificação binária** (binary classification), onde temos apenas duas classes disponíveis.

Model Evaluation

- Em nosso exemplo, tentaremos prever se uma imagem é um cachorro ou um gato.
- Como isso é aprendizado supervisionado, primeiro **ajustaremos/treinaremos um modelo** (fit/train) nos **dados de treinamento** (training data) e, em seguida, **testaremos** (test) o modelo nos **dados de teste** (testing data).
- Assim que obtivermos as previsões do modelo a partir dos dados **X_test**, compararemos com os **valores y verdadeiros** (os rótulos corretos).

Model Evaluation



TRAINED
MODEL

Model Evaluation



**Test Image
from X_{test}**

TRAINED
MODEL

Model Evaluation



**Test Image
from X_{test}**

DOG

**Correct Label
from y_{test}**

TRAINED
MODEL

Model Evaluation



**Test Image
from X_{test}**

DOG

**Correct Label
from y_{test}**

TRAINED
MODEL

DOG

**Prediction on
Test Image**

Model Evaluation



Test Image
from X_{test}


TRAINED
MODEL

DOG

Prediction on
Test Image

DOG

Correct Label
from y_{test}

 DOG == DOG ?

Compare Prediction to Correct Label

Model Evaluation



**Test Image
from X_{test}**

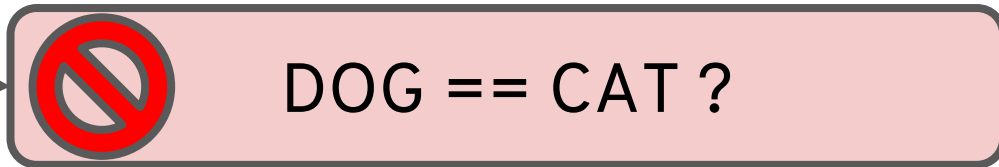


CAT

**Prediction on
Test Image**

DOG

**Correct Label
from y_{test}**



Compare Prediction to Correct Label

Model Evaluation

- Repetimos esse processo para todas as imagens em nossos dados X_{test} .
- No final teremos uma contagem de correspondências corretas e uma contagem de correspondências incorretas.
- A principal constatação que precisamos fazer é que: **no mundo real, nem todas as correspondências incorretas ou corretas têm o mesmo valor!**

Model Evaluation

- Também, no mundo real, uma única métrica não contará a história completa!
- Para entender tudo isso, vamos trazer de volta as 4 métricas que mencionamos e ver como elas são calculadas.
- Poderíamos organizar nossos valores previstos em comparação com os valores reais em uma **matriz de confusão** (confusion matrix).

Model Evaluation

- Accuracy
 - Acurácia, em problemas de classificação, **é o número de previsões corretas** feitas pelo modelo dividido pelo **número total de previsões**.

Model Evaluation

- Accuracy
 - Por exemplo, se o conjunto X_test tiver 100 imagens e nosso modelo previu **corretamente** 80 imagens, teremos **80/100**.
 - **0.8** ou **80% accuracy**.

Model Evaluation

- Accuracy
 - Acurácia é útil quando as classes de destino são bem equilibradas
 - Em nosso exemplo, teríamos aproximadamente a mesma quantidade de imagens de gatos e cães.

Model Evaluation

- Accuracy

- Acurácia **não** é uma boa escolha com classes **desequilibradas!**
- Imagine que tivéssemos 99 imagens de cães e 1 imagem de um gato.
- Se nosso modelo fosse simplesmente uma linha que sempre previsse **cachorro**, obteríamos 99% de acurácia!

Model Evaluation

- Accuracy

- Acurácia **não** é uma boa escolha com classes **desequilibradas!**
- Imagine que tivéssemos 99 imagens de cães e 1 imagem de um gato.
- Se nosso modelo fosse simplesmente uma linha que sempre previsse **cachorro**, obteríamos 99% de acurácia!
- Nesta situação, é útil entender a definição de **recall** e **precision**

Model Evaluation

- Recall
 - Capacidade de um modelo para encontrar todos os casos relevantes dentro de um conjunto de dados.
 - A definição precisa de revocação é: o número de verdadeiros positivos (true positives) **dividido** pelo número de verdadeiros positivos mais o número de falsos negativos (false negatives).

Model Evaluation

- Precision

- Capacidade de um modelo de classificação para identificar apenas os pontos de dados relevantes.
- A precisão é definida como o número de verdadeiros positivos (true positives) **dividido** pelo número de verdadeiros positivos mais o número de falsos positivos (false positives).

Model Evaluation

- Recall e Precision
 - Muitas vezes você tem um trade-off entre Revocação e Precisão.
 - Enquanto a revocação expressa a capacidade de encontrar todas as instâncias relevantes em um conjunto de dados, a precisão expressa a proporção dos pontos de dados que nosso modelo diz que são relevantes, são de fato relevantes.

Model Evaluation

- F1-Score
 - Nos casos em que queremos encontrar uma combinação ideal de precision e recall, podemos combinar as duas métricas usando o que é chamado de Pontuação-F1.

Model Evaluation

- F1-Score

- A pontuação F1 é a média harmônica de precisão e recall levando ambas as métricas em consideração na equação a seguir:

$$F_1 = 2 * \frac{precision * recall}{precision + recall}$$

Model Evaluation

- F1-Score
 - Utilizamos a média harmônica em vez de uma média simples porque ela pune valores extremos.
 - Um classificador com precisão de 1.0 e recall de 0.0 tem uma média simples de 0.5, mas uma pontuação F1 de 0.

Model Evaluation

- Também podemos visualizar todas as imagens classificadas corretamente versus classificadas incorretamente na forma de uma matriz de confusão (confusion matrix).

Confusion Matrix

		predicted condition	
total population		prediction positive	prediction negative
true condition	condition positive	True Positive (TP)	False Negative (FN) (type II error)
	condition negative	False Positive (FP) (Type I error)	True Negative (TN)

		predicted condition		
total population		prediction positive	prediction negative	Prevalence $= \frac{\Sigma \text{ condition positive}}{\Sigma \text{ total population}}$
true condition	condition positive	True Positive (TP)	False Negative (FN) (type II error)	True Positive Rate (TPR), Sensitivity, Recall, Probability of Detection $= \frac{\Sigma \text{ TP}}{\Sigma \text{ condition positive}}$
	condition negative	False Positive (FP) (Type I error)	True Negative (TN)	False Positive Rate (FPR), Fall-out, Probability of False Alarm $= \frac{\Sigma \text{ FP}}{\Sigma \text{ condition negative}}$
Accuracy $= \frac{\Sigma \text{ TP} + \Sigma \text{ TN}}{\Sigma \text{ total population}}$		Positive Predictive Value (PPV), Precision $= \frac{\Sigma \text{ TP}}{\Sigma \text{ prediction positive}}$	False Omission Rate (FOR) $= \frac{\Sigma \text{ FN}}{\Sigma \text{ prediction negative}}$	Positive Likelihood Ratio (LR+) $= \frac{\text{TPR}}{\text{FPR}}$
		False Discovery Rate (FDR) $= \frac{\Sigma \text{ FP}}{\Sigma \text{ prediction positive}}$	Negative Predictive Value (NPV) $= \frac{\Sigma \text{ TN}}{\Sigma \text{ prediction negative}}$	Negative Likelihood Ratio (LR-) $= \frac{\text{FNR}}{\text{TNR}}$

Model Evaluation

- O ponto principal a ser lembrado com a matriz de confusão (confusion matrix) e as várias métricas calculadas é que todas elas são fundamentalmente maneiras de comparar os valores previstos versus os valores verdadeiros.
- O que constitui métricas “boas” dependerá realmente da situação específica!

Model Evaluation

- Ainda confuso sobre a matriz de confusão (confusion matrix)?
- Sem problemas! Confira a página da Wikipedia, ela tem um diagrama muito bom com todas as fórmulas para todas as métricas.
- Ao longo do treinamento, geralmente apenas imprimimos métricas (por exemplo, acurácia).

Model Evaluation

- Vamos relembrar essa ideia de:
 - O que é uma acurácia boa o suficiente?
- Tudo depende do contexto da situação!
- Você criou um modelo para prever a presença de uma doença?
- A presença da doença é bem equilibrada na população em geral? (Provavelmente não!)

Model Evaluation

- Muitas vezes, os modelos são usados como testes de diagnóstico rápidos **antes** de fazer um teste mais invasivo (por exemplo, fazer um teste de urina antes de fazer uma biópsia)
- Também precisamos considerar o que está em jogo!

Model Evaluation

- Muitas vezes temos um trade-off entre precision/recall. Precisamos decidir se o modelo deve se concentrar na correção de falsos positivos (false positives) versus falsos negativos (false negatives).
- No diagnóstico de doenças, provavelmente é melhor ir na direção dos falsos positivos, por isso nos certificamos de classificar corretamente o maior número possível de casos de doenças!

Model Evaluation

- Tudo isso para dizer que machine learning não é realizado em um “vácuo”, mas sim em um processo colaborativo onde devemos consultar especialistas no domínio (por exemplo, médicos)

Avaliando o Desempenho (Evaluating Performance)

REGRESSÃO
(REGRESSION)

Evaluating Regression

- Vamos tomar um momento agora para discutir a avaliação de modelos de regressão
- A regressão é a tarefa executada por um modelo quando ele tenta prever valores contínuos (ao contrário de valores categóricos, que é classificação)

Evaluating Regression

- Já discutimos algumas métricas de avaliação, como accuracy ou recall.
- Esses tipos de métrica não são úteis para problemas de regressão, precisamos de métricas projetadas para valores contínuos!

Evaluating Regression

- Por exemplo, tentar prever o preço de uma casa com base em suas características é uma **tarefa de regressão**.
- Tentar prever o país em que uma casa se encontra, considerando suas características, seria uma tarefa de classificação.

Evaluating Regression

- Vamos discutir algumas das métricas de avaliação mais comuns para regressão:
 - Erro Absoluto Médio (Mean Absolute Error)
 - Erro Quadrático Médio (Mean Squared Error)
 - Distância Quadrática Média (Root Mean Square Error)

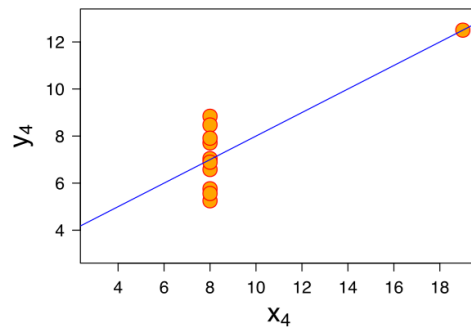
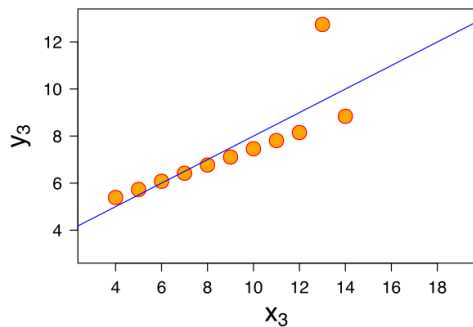
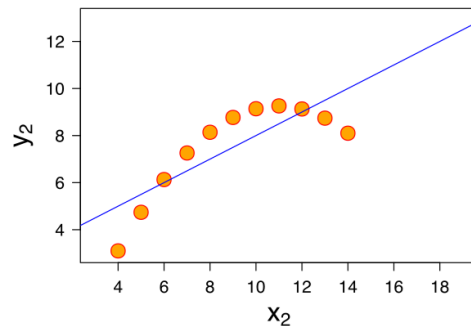
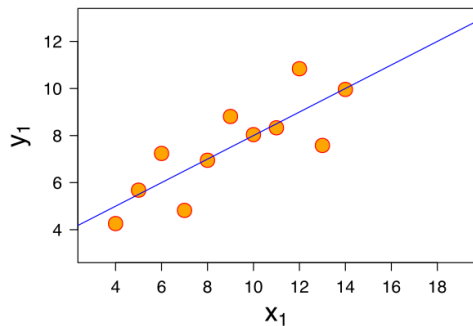
Evaluating Regression

- Erro Absoluto Médio (**Mean Absolute Error - MAE**)
 - É a média do valor absoluto dos erros.
 - Fácil de entender

$$\frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

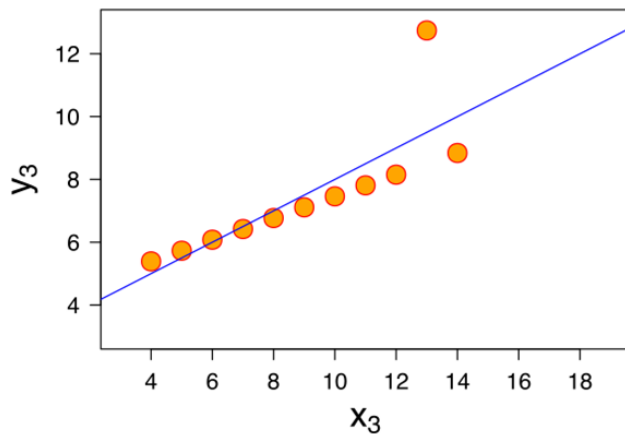
Evaluating Regression

- O MAE, no entanto, não punirá grandes erros.



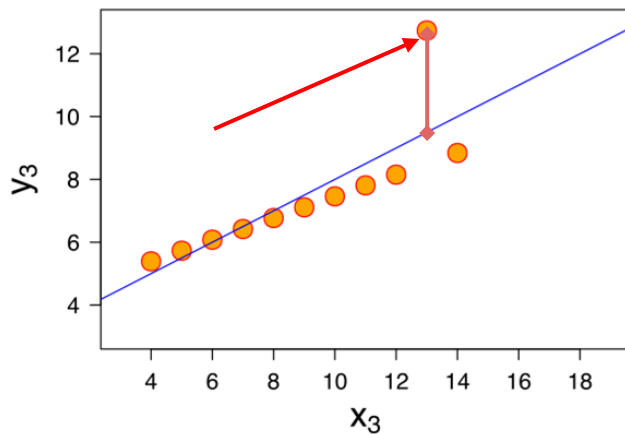
Evaluating Regression

- O MAE, no entanto, não punirá grandes erros.



Evaluating Regression

- Queremos que nossas métricas de erro levem isso em consideração!



Evaluating Regression

- Erro Quadrático Médio (**Mean Squared Error - MSE**)
 - É a média dos erros ao quadrado.
 - Erros maiores são mais perceptíveis em comparação ao MAE, o que torna o MSE um método mais popular que o MAE.

$$\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

Evaluating Regression

- Distância Quadrática Média (Root Mean Square Error - RMSE)
 - É a raiz quadrada da média dos erros ao quadrado.
 - É o mais popular (tem as mesmas unidades que y)

$$\sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

Machine Learning

- Dúvida mais comuns dos alunos:
 - “Este valor de RMSE é bom?”
- Contexto é tudo!
- Um RMSE de R\$ 10 é fantástico para prever o preço de uma casa, mas horrível para prever o preço de uma barra de chocolate!

Machine Learning

- Compare sua métrica de erro com o valor médio do rótulo (label) em seu conjunto de dados para tentar obter uma intuição de seu desempenho geral.
- O conhecimento do domínio sendo estudado também desempenha um papel importante aqui!

Machine Learning

- Também é necessário considerar o contexto de importância em que será aplicado o modelo.
- Por exemplo, podemos criar um modelo para prever quanta medicação deve ser administrada a um paciente. Neste caso, pequenas flutuações no RMSE podem ser muito significativas.

Aprendizado Não Supervisionado (Unsupervised Learning)

Machine Learning

- Vimos o aprendizado supervisionado, onde o **rótulo era conhecido** devido ao **histórico de dados rotulados** (historical labeled data).
- Mas o que acontece quando não temos um histórico de rótulos?

Machine Learning

- Existem certas tarefas que se enquadram no aprendizado não supervisionado:
 - Agrupamento (Clustering)
 - Detecção de Anomalia (Anomaly Detection)
 - Redução de Dimensionalidade (Dimensionality Reduction)

Machine Learning

- Clustering
 - Agrupando pontos de **dados não rotulados** (unlabeled data) em categorias/clusters.
 - Os *data points* são atribuídos a um *cluster* com base em sua semelhança.

Machine Learning

- Anomaly Detection
 - Tentativas de detectar discrepâncias em um conjunto de dados.
 - Por exemplo, transações fraudulentas em um cartão de crédito.

Machine Learning

- Dimensionality Reduction
 - Técnicas de processamento de dados que reduzem o número de *features* em um conjunto de dados, seja para compactação ou para entender melhor as tendências implícitas em um conjunto de dados.

Machine Learning

- Unsupervised Learning
 - É importante observar que essas são situações em que **não temos** a resposta correta para os dados históricos!
 - O que significa que a avaliação é muito mais difícil e com mais nuances!

Unsupervised Process



Machine Learning

- Futuramente exploraremos processos de aprendizado não supervisionados com estruturas de rede neural especializadas, como *autoencoders*.