

dnc

apresenta



Data Science & Machine Learning



Data Science & Machine Learning

Bem vindos a
segunda parte do
módulo

Aula 12: ML Classificação

Consultor: Daniel Soria

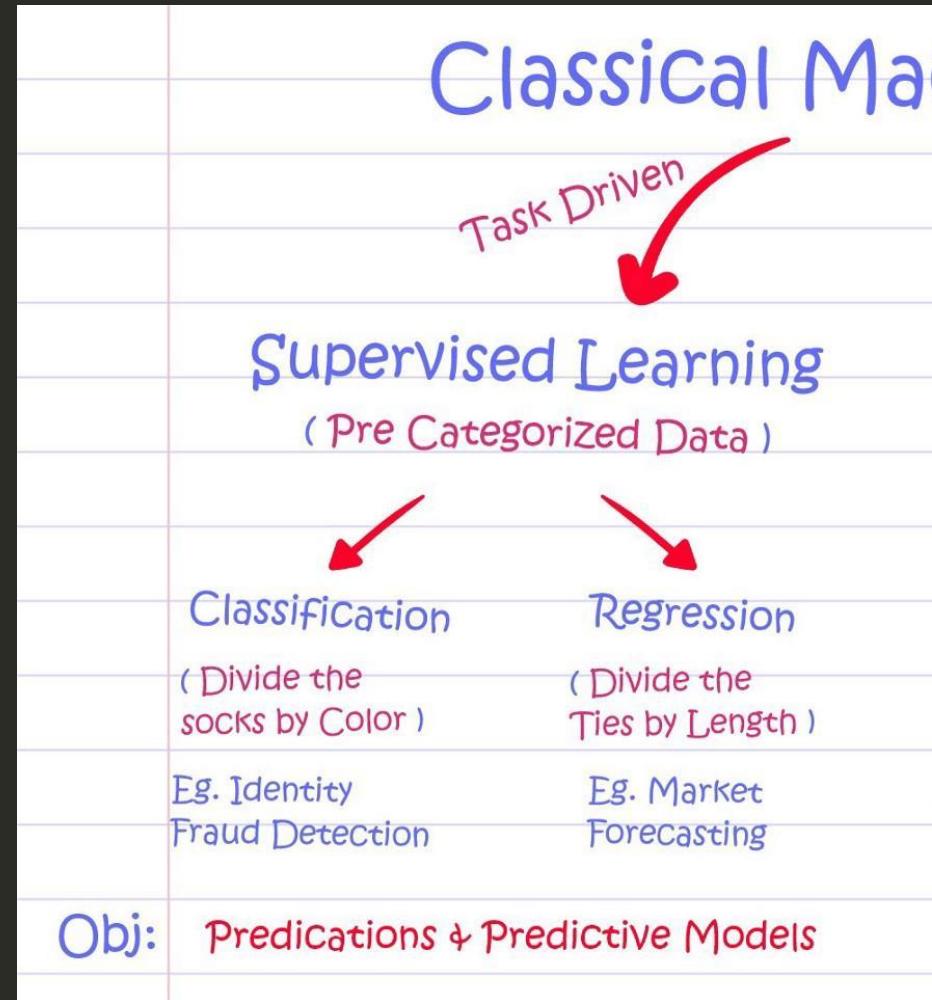


Classificação são modelos que tentam rotular um valor em uma classe de valores similares

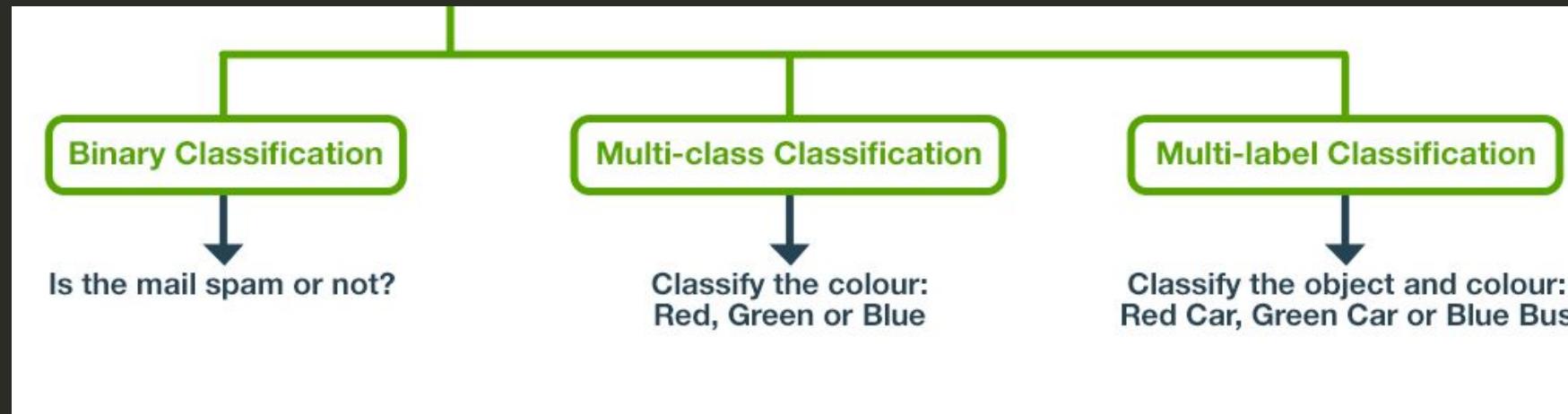
Não Ácero x Áceros



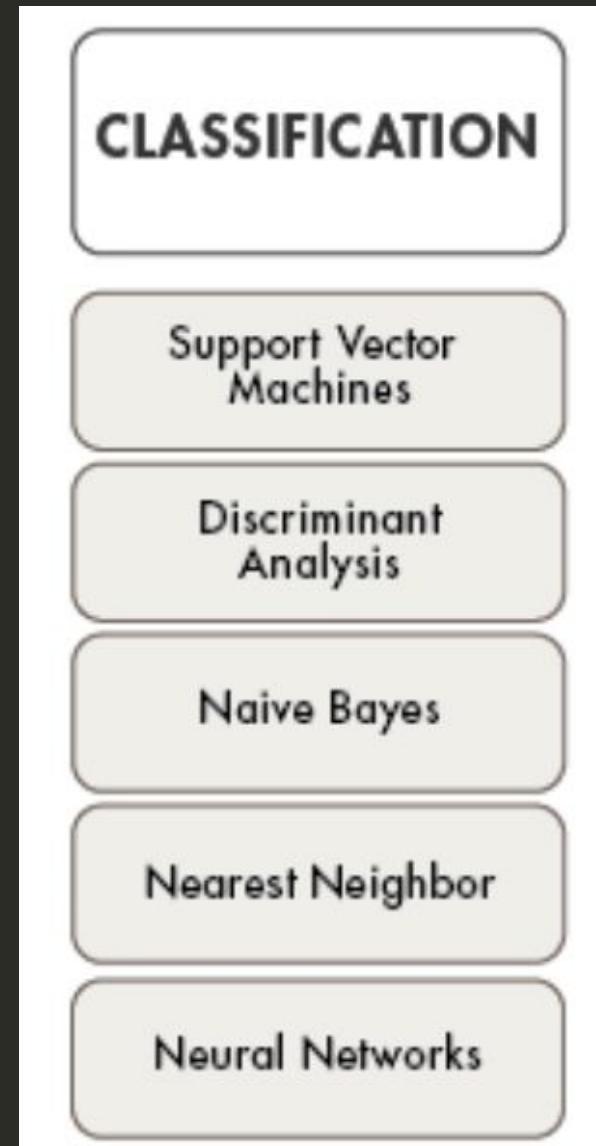
Classificação



Classificação

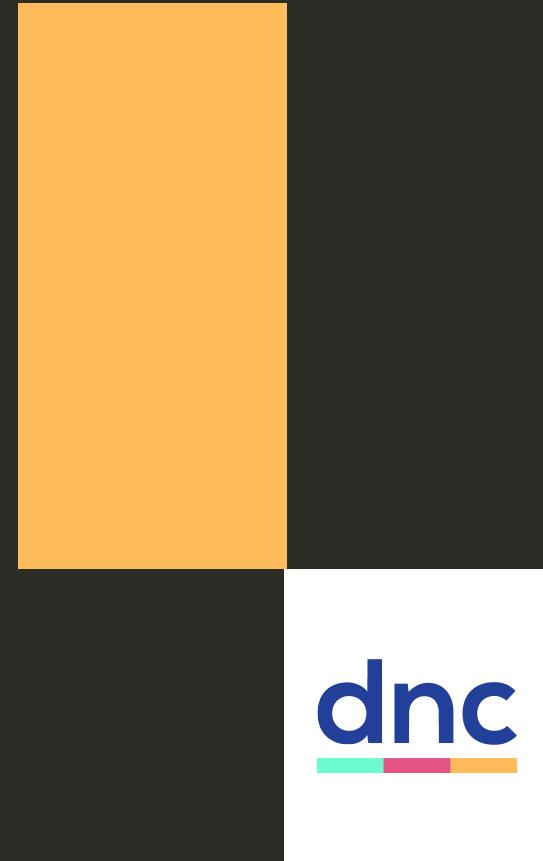


Classificação





Data Science & Machine Learning

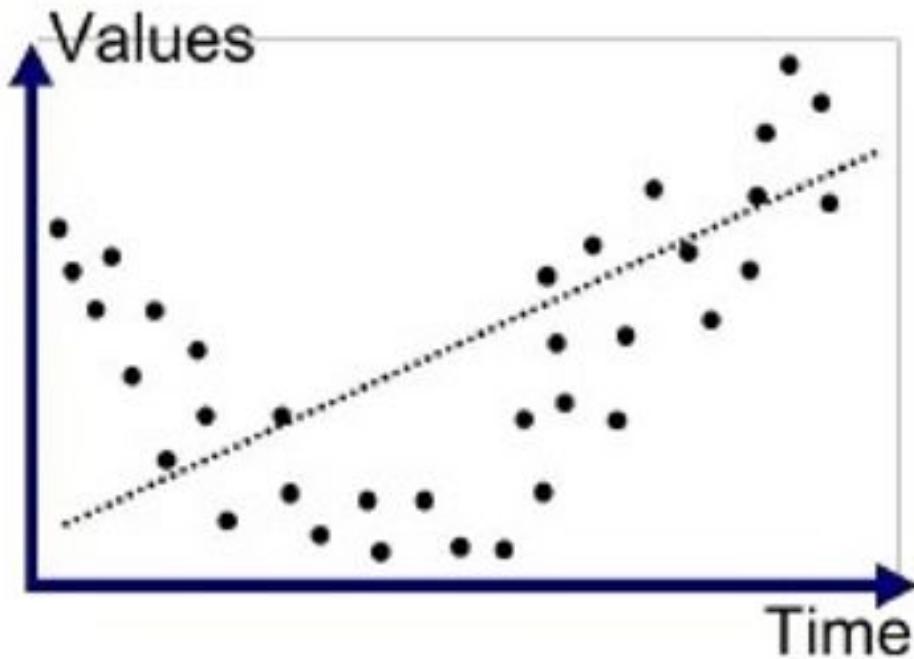


Aula 14: Overfitting e Underfitting

Underfitting é quando o modelo ficou MUITO frouxo, Overfitting é quando o modelo ficou MUITO ajustado

Underfitting ocorre quando o teu modelo está performando muito mal em cima da tua massa de teste.

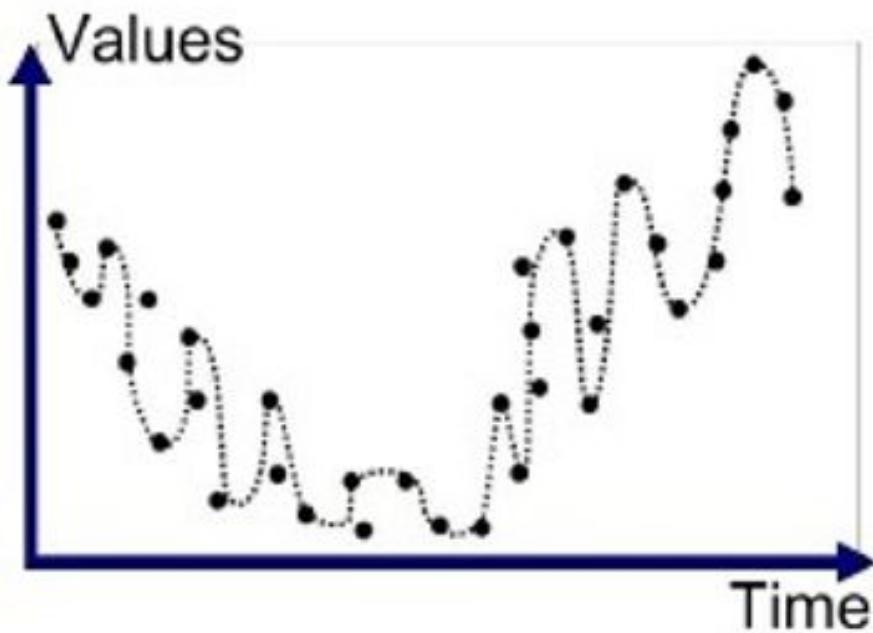
Você provavelmente precisará mudar de modelo.



Underfitted

Overfitting ocorreu quando o teu modelo está performando muito bem em cima da tua massa de teste.

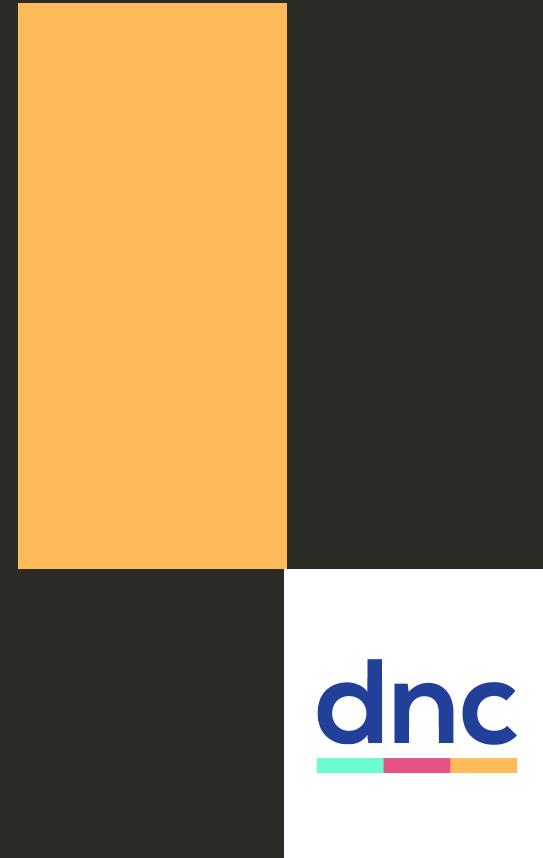
Porém quando vai na vida real o modelo não acerta pois está muito ajustado ao que foi treinado.



Overfitted



Data Science & Machine Learning



Aula 13: Métricas de avaliação de modelos Supervisionados

Train Data Set

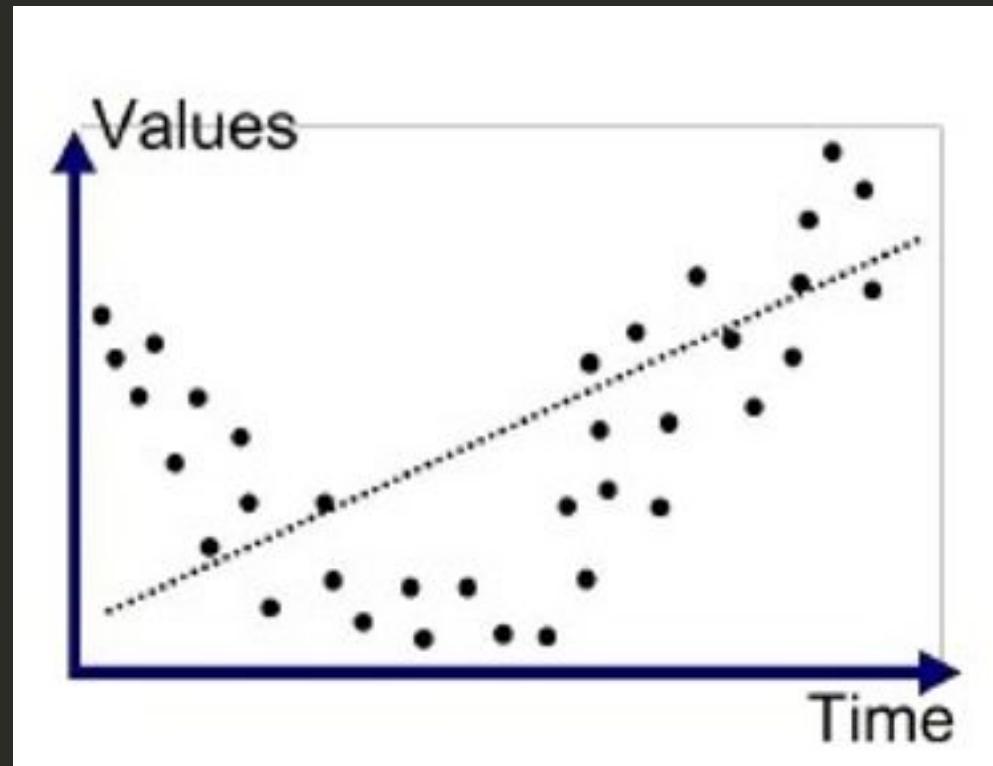
X

Test Data Set

Métricas de Avaliação de Regressão

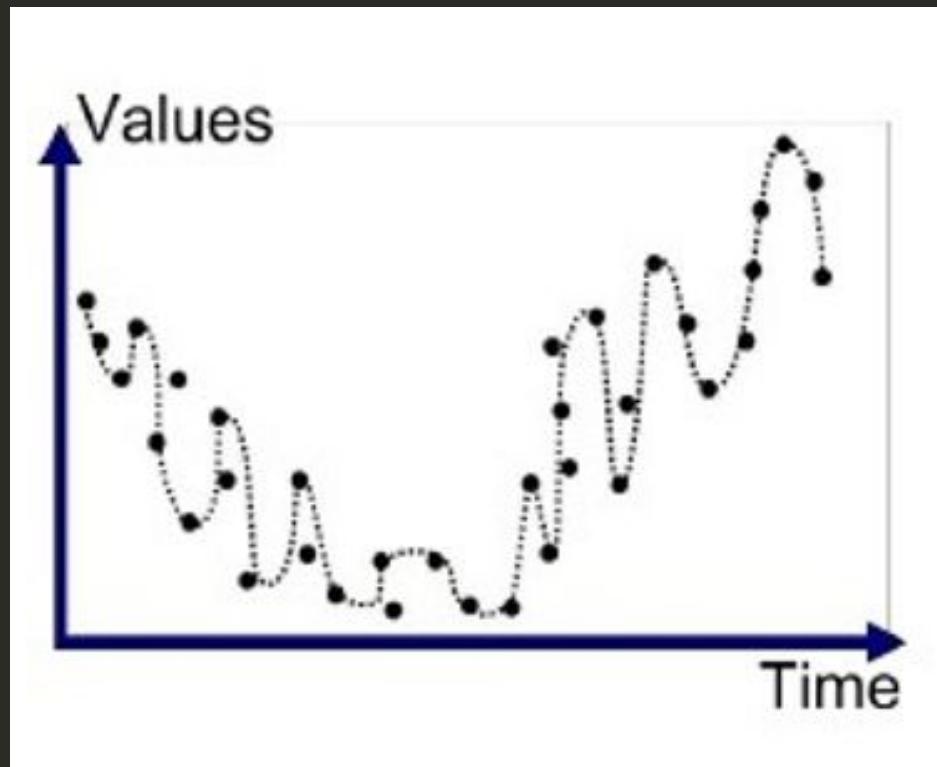
- **R Square/Adjusted R Square**
- **Mean Square Error(MSE)/Root Mean Square Error(RMSE)**
- **Mean Absolute Error(MAE)**

- R Square/Adjusted R Square



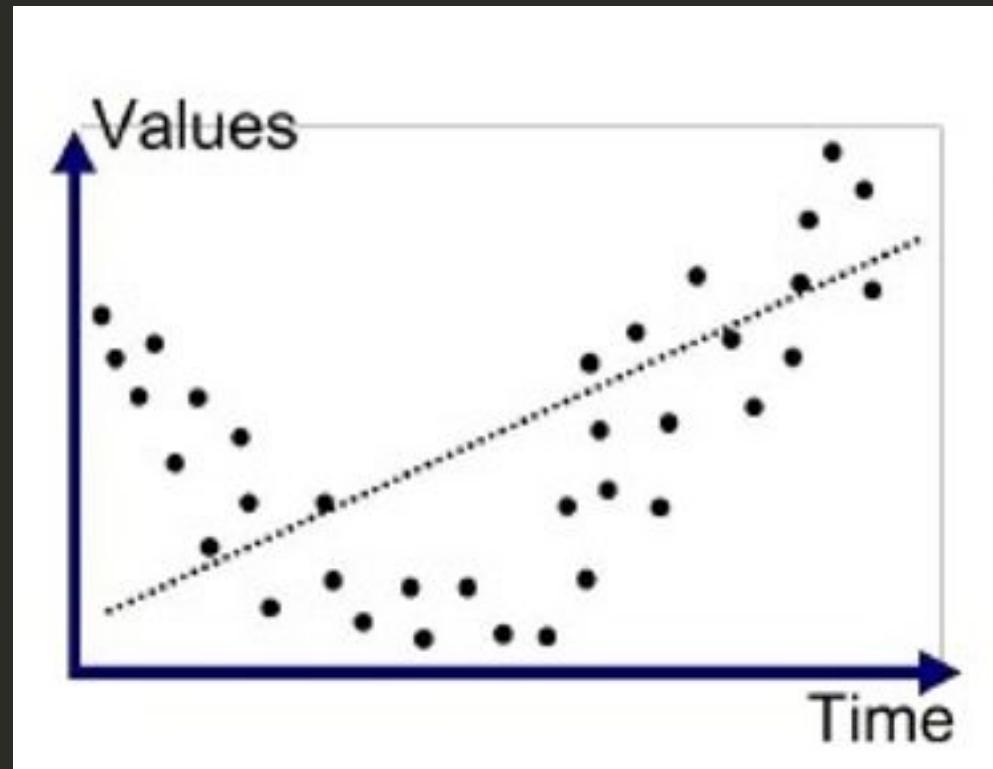
R Square = 0.01

- R Square/Adjusted R Square

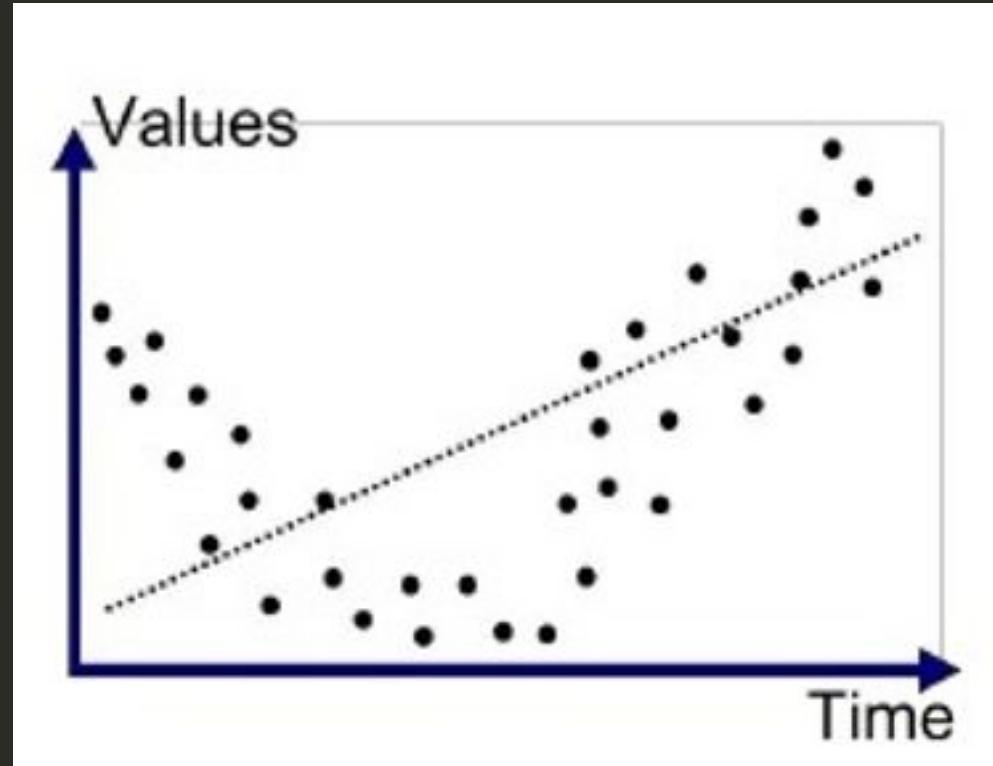


R Square = 1.0

- MSE/ RMSE



- MAE – Mean Absolut Error



Métricas de Avaliação de Regressão

- **R Square/Adjusted R Square**
- **Mean Square Error(MSE)/Root Mean Square Error(RMSE)**
- **Mean Absolute Error(MAE)**

Métricas de Avaliação de Classificação

- Acuracidade
- Precisão
- Recall
- F1 Score

Matrix de Confusão



Classificado

Real

	NÃO É GATOS	É GATO
NÃO É GATO	99	1
É GATO	0	0

Acuracidade = 99%

Matrix de Confusão



Classificado

		NÃO É GATOS	É GATO
Real	NÃO É GATO	98	1
	É GATO	1	0

Acuracidade = 98%

Matrix de Confusão



Classificado

Real

	NEGATIVO	POSITIVO
NEGATIVO		
POSITIVO		

Matrix de Confusão



Classificado

Real

	NEGATIVO	POSITIVO
NEGATIVO	VERDADEIRO NEGATIVO	FALSO POSITIVO
POSITIVO	FALSO NEGATIVO	VERDADEIRO POSITIVO

$$Accuracy = \frac{TruePositive + TrueNegative}{TotalSample}$$

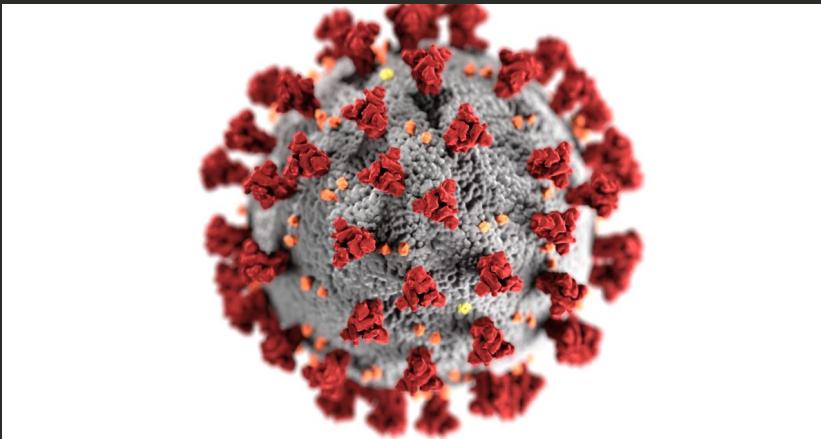
Matrix de Confusão



		Classificado	
		NEGATIVO	POSITIVO
Real	NEGATIVO	VERDADEIRO NEGATIVO	FALSO POSITIVO
	POSITIVO	FALSO NEGATIVO	VERDADEIRO POSITIVO

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} = \frac{\text{True Positive}}{\text{Total Predicted Positive}}$$

Matrix de Confusão



		Classificado	
		NEGATIVO	POSITIVO
Real	NEGATIVO	VERDADEIRO NEGATIVO	FALSO POSITIVO
	POSITIVO	FALSO NEGATIVO	VERDADEIRO POSITIVO

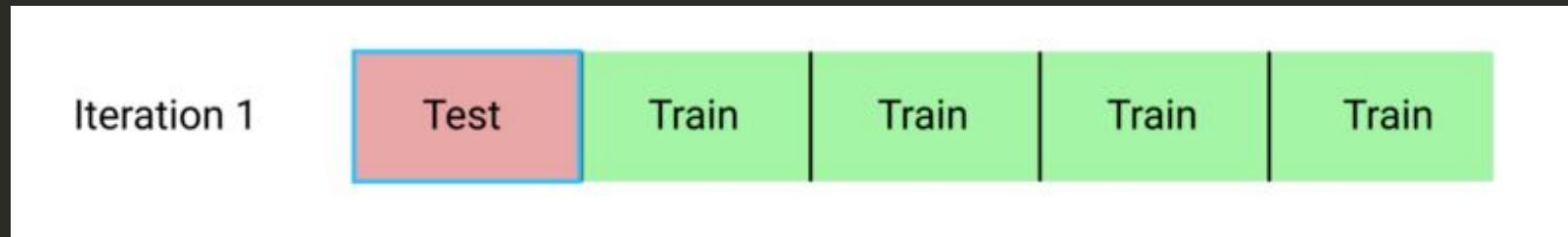
$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} = \frac{\text{True Positive}}{\text{Total Actual Positive}}$$

Matrix de Confusão

$$\text{Accuracy} = \frac{\text{TruePositive} + \text{TrueNegative}}{\text{TotalSample}}$$

$$\text{F1} = 2 \times \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}$$

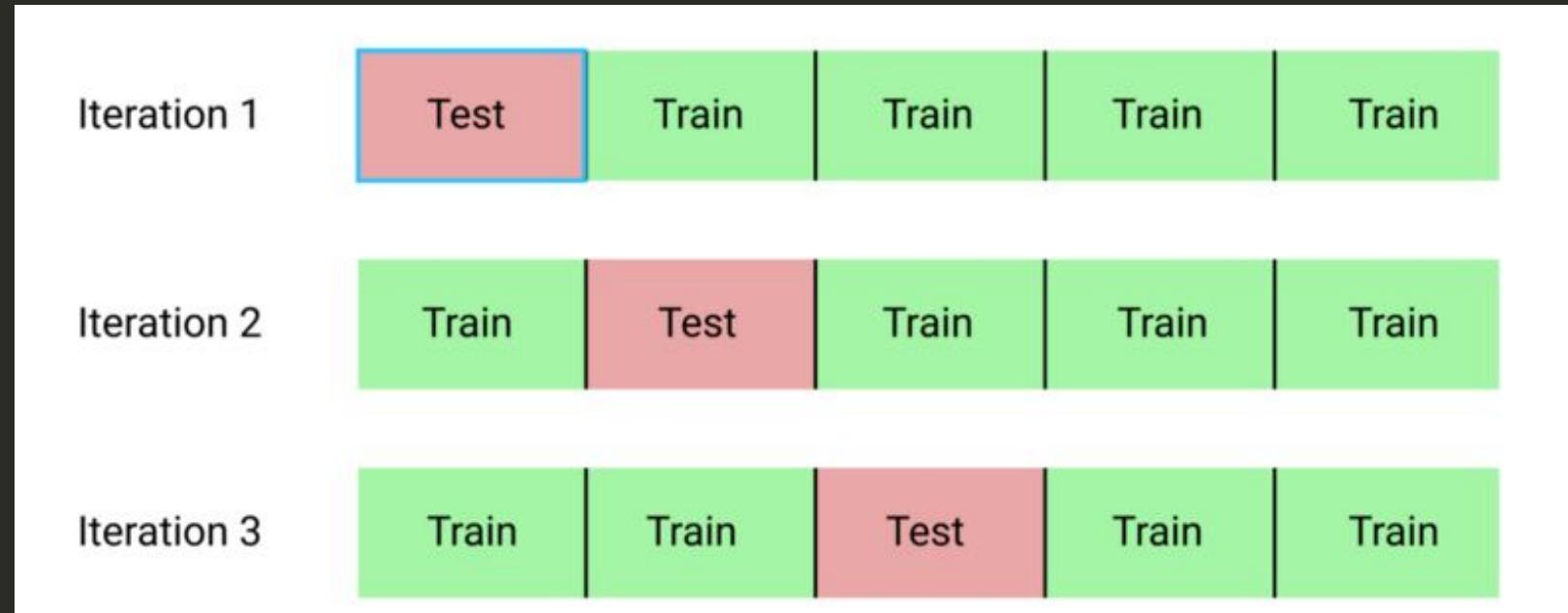
Cross Validation



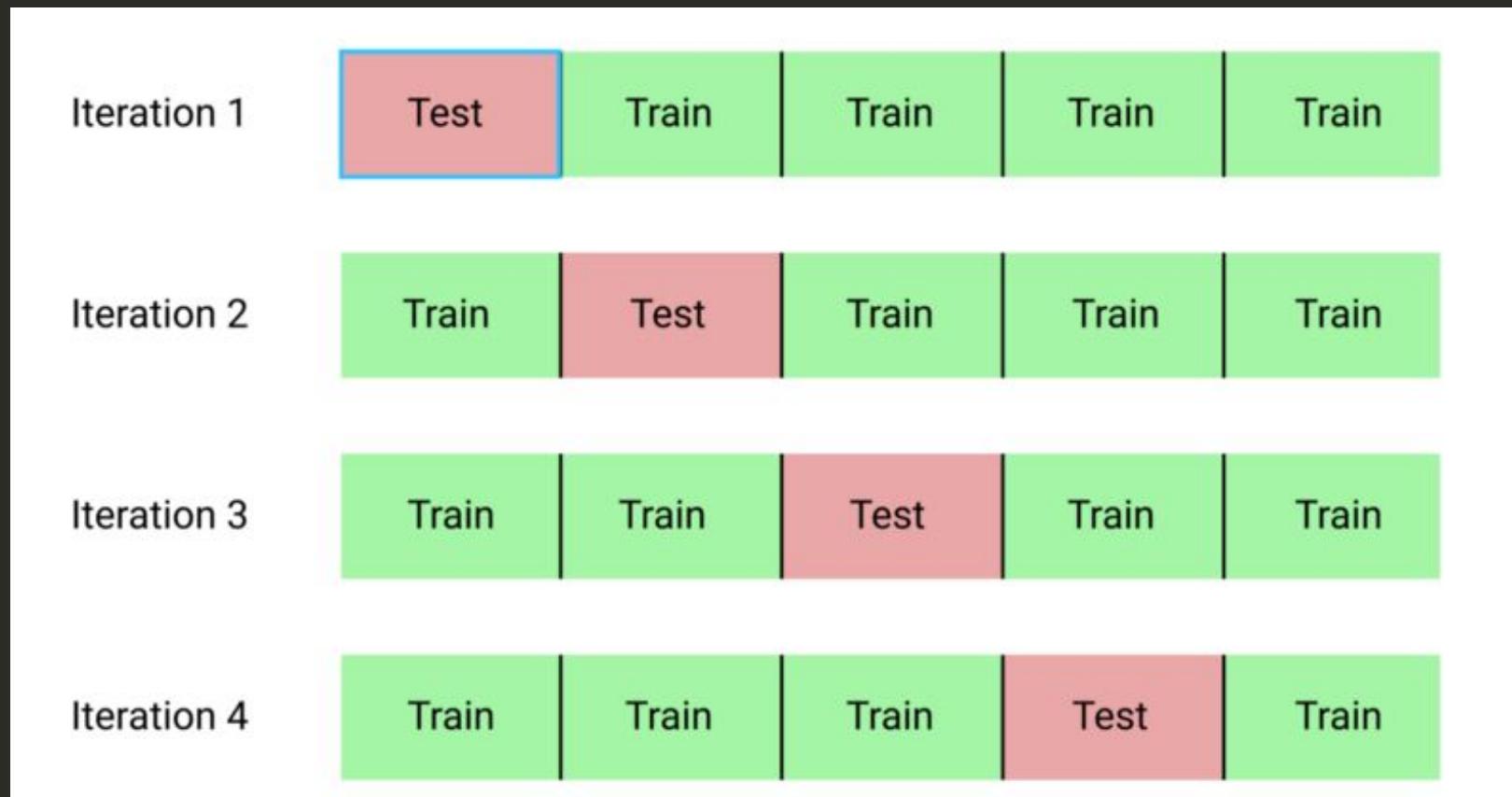
Cross Validation



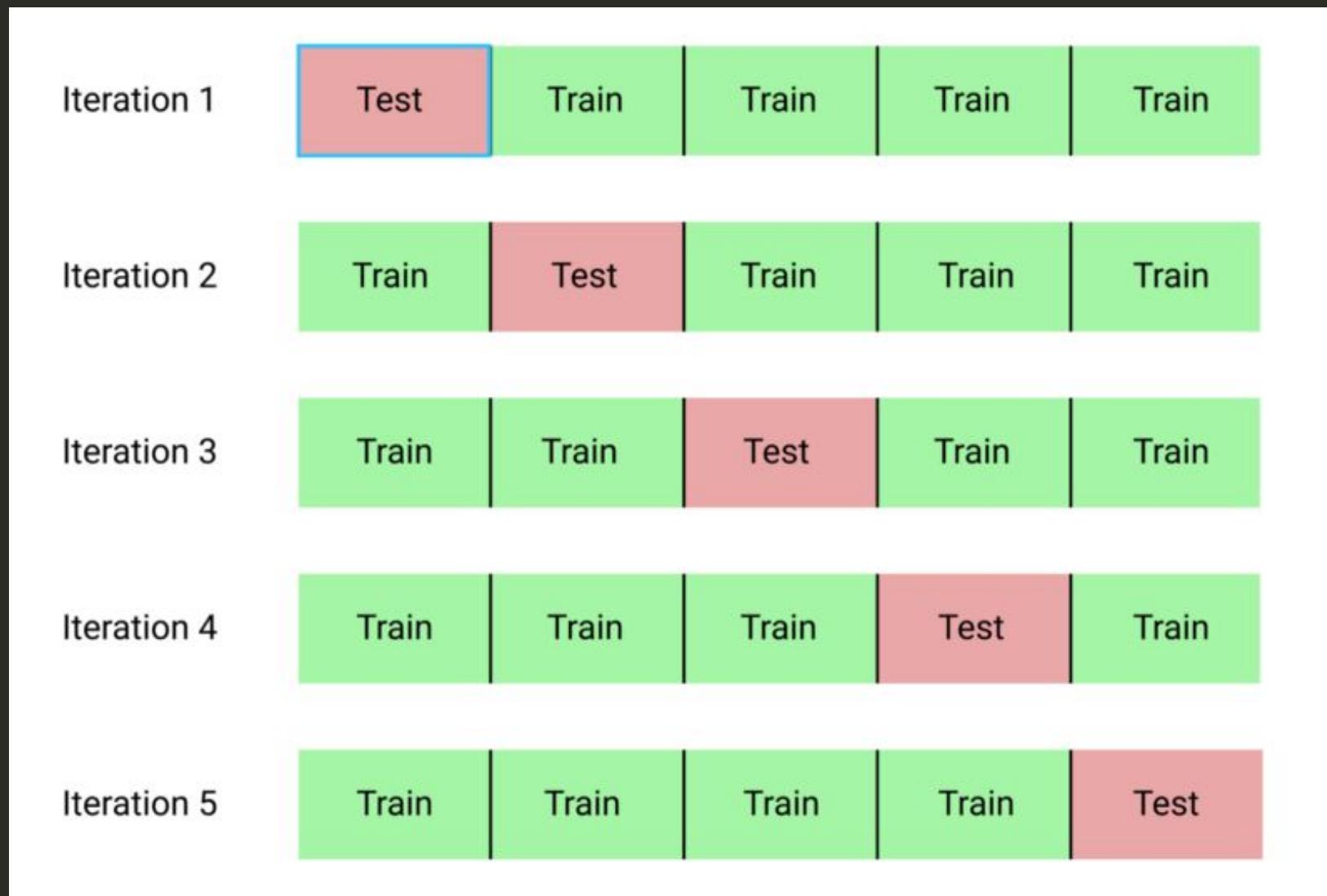
Cross Validation



Cross Validation

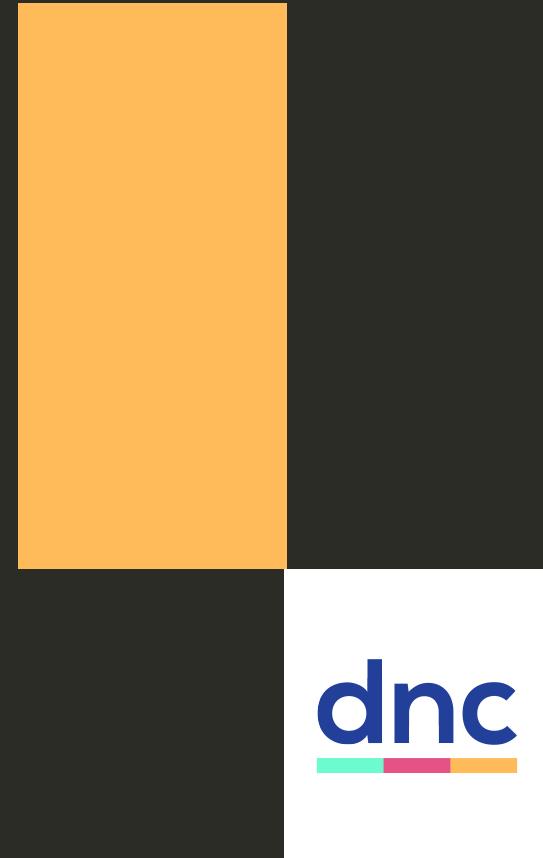


K-Fold Cross Validation





Data Science & Machine Learning



Aula 15: Machine Learning Não supervisionado



supervised learning

Input data



Annotations

These are
apples



Model

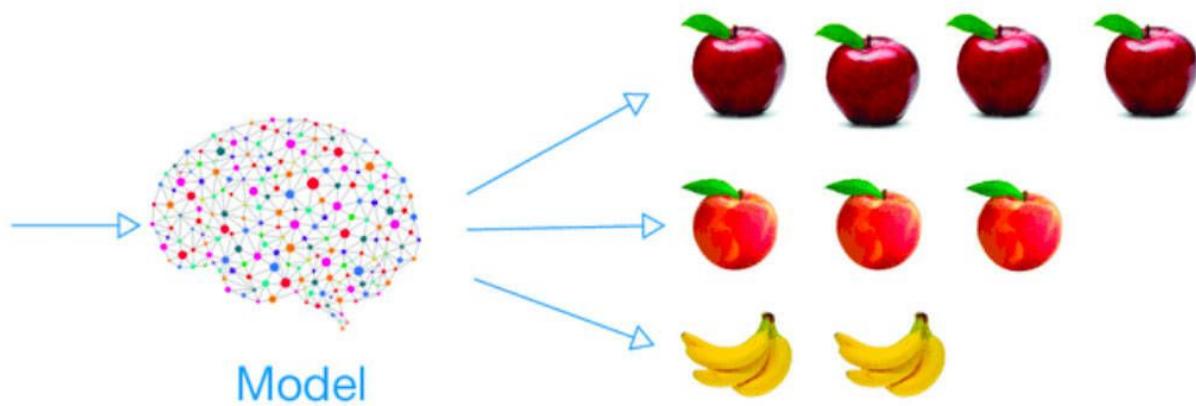


Prediction

Its an
apple!

unsupervised learning

Input data



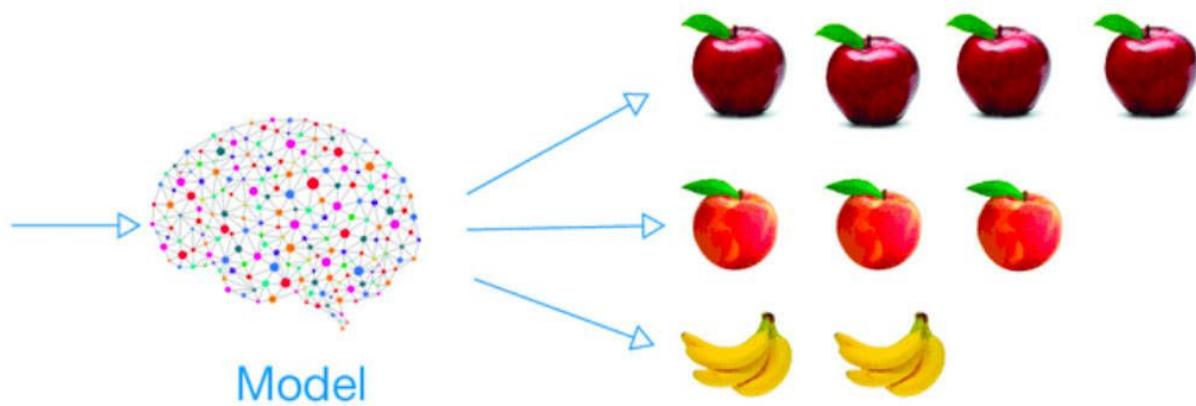
DETECTAR ANOLAMIAS

Foi uma Fraude ou não?

INPUT	PROCESS	OUTPUT
x1 = Valor da transação	Função para Detectar uma fraude no cartão de crédito	Transação foi fraude ou não
x2 = Site da compra		
x3 = Horário da compra		
x4 = Movimentação da conta		

unsupervised learning

Input data

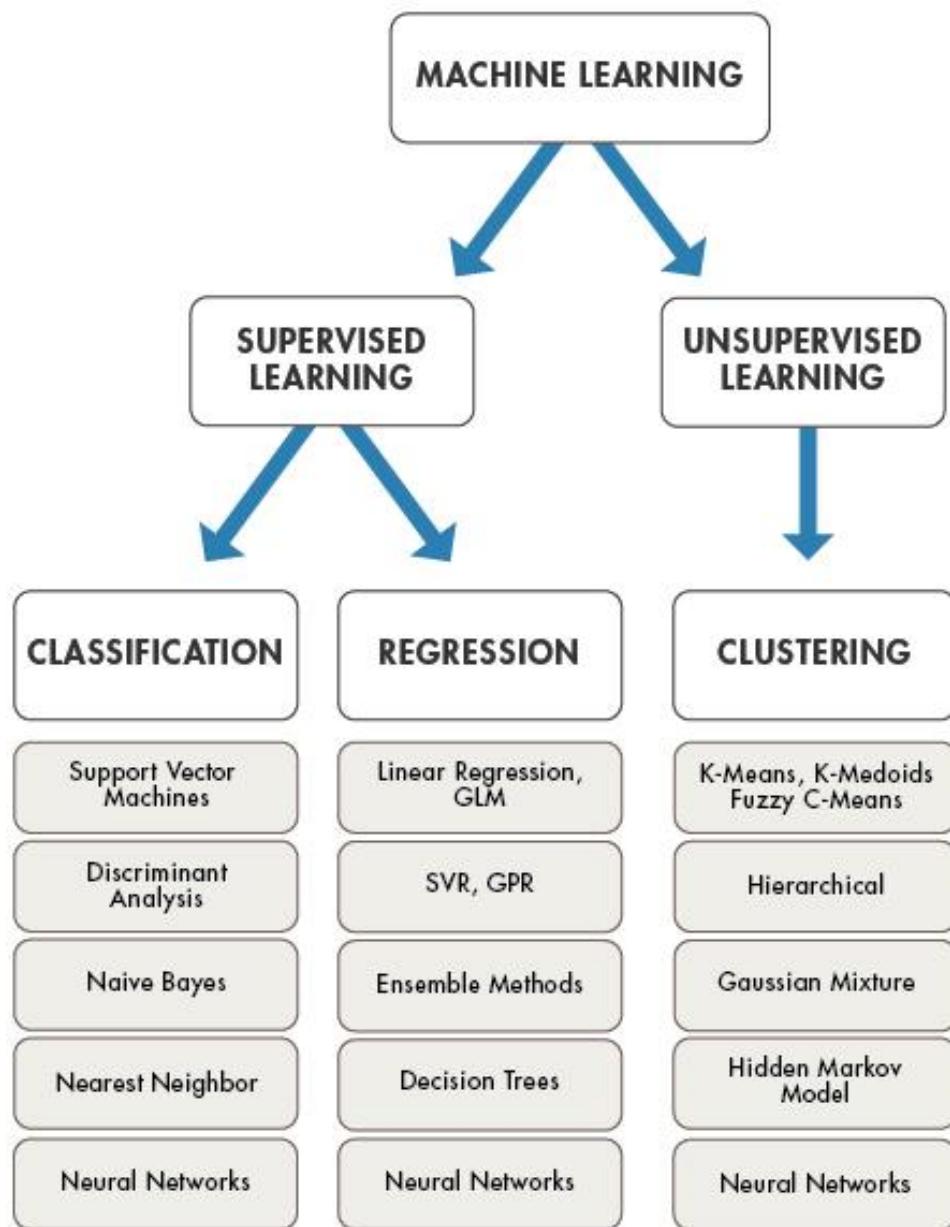


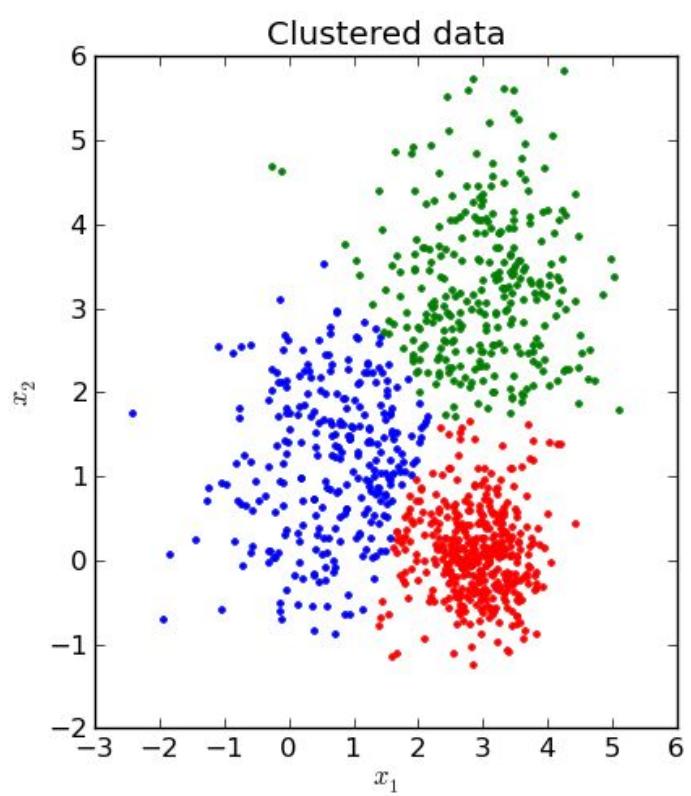
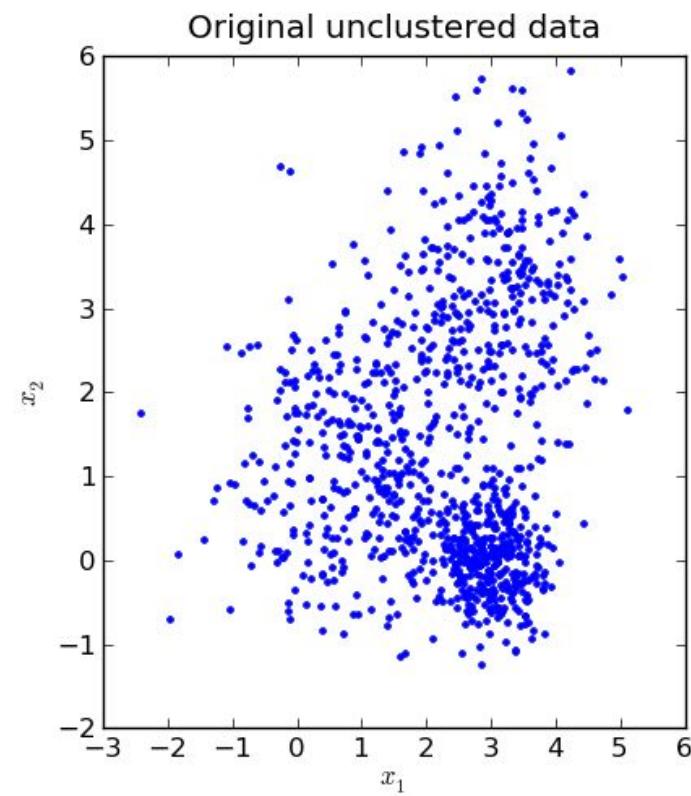


Data Science & Machine Learning

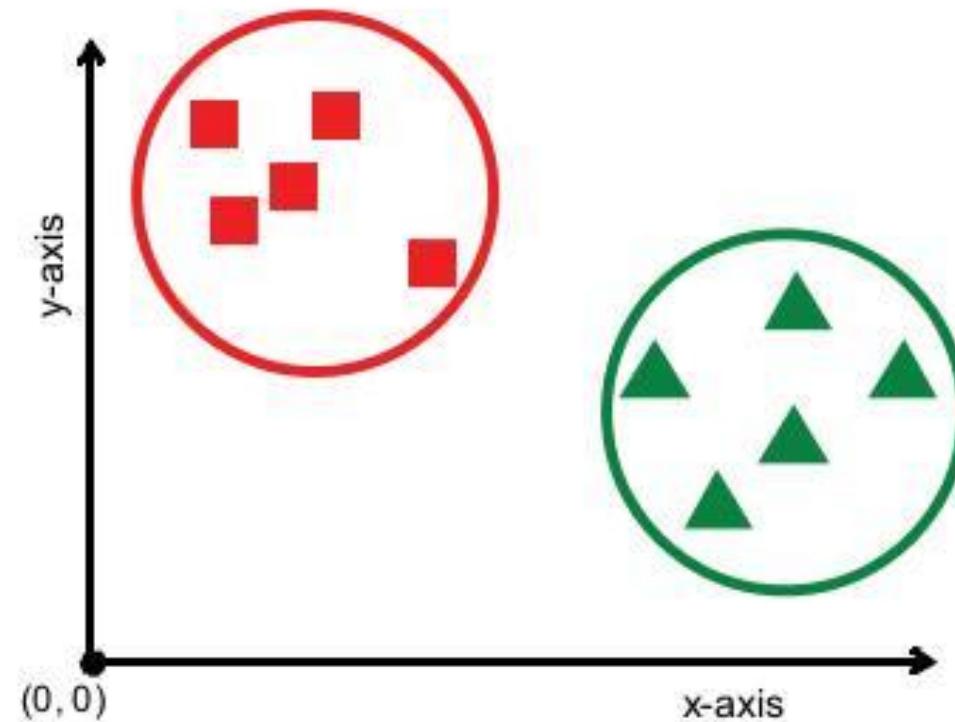


Aula 16: ML Clustering

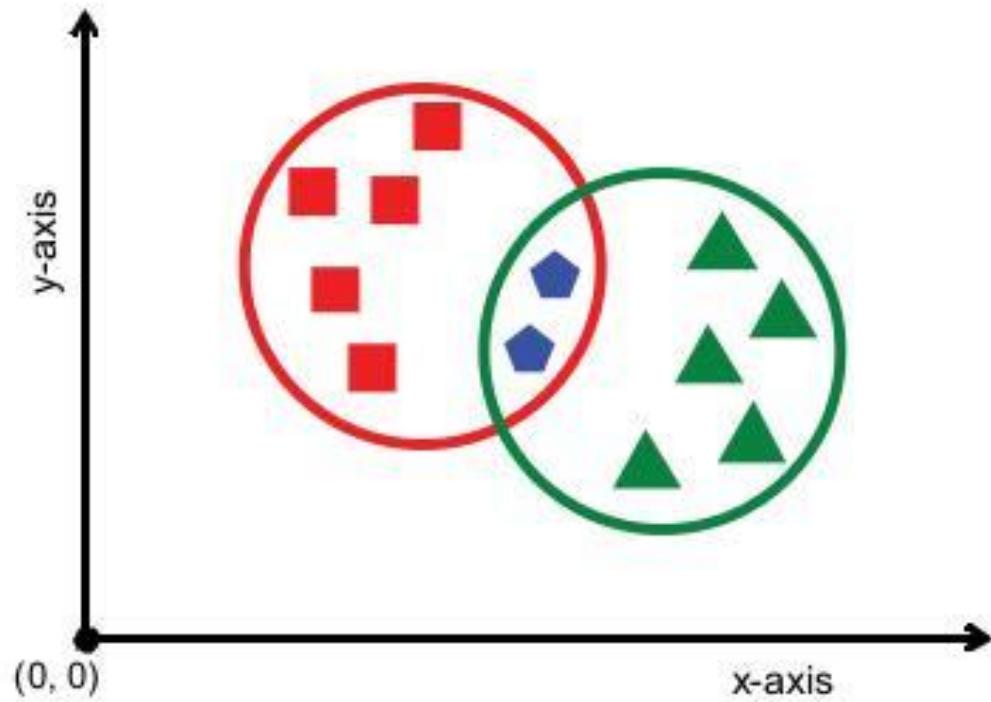


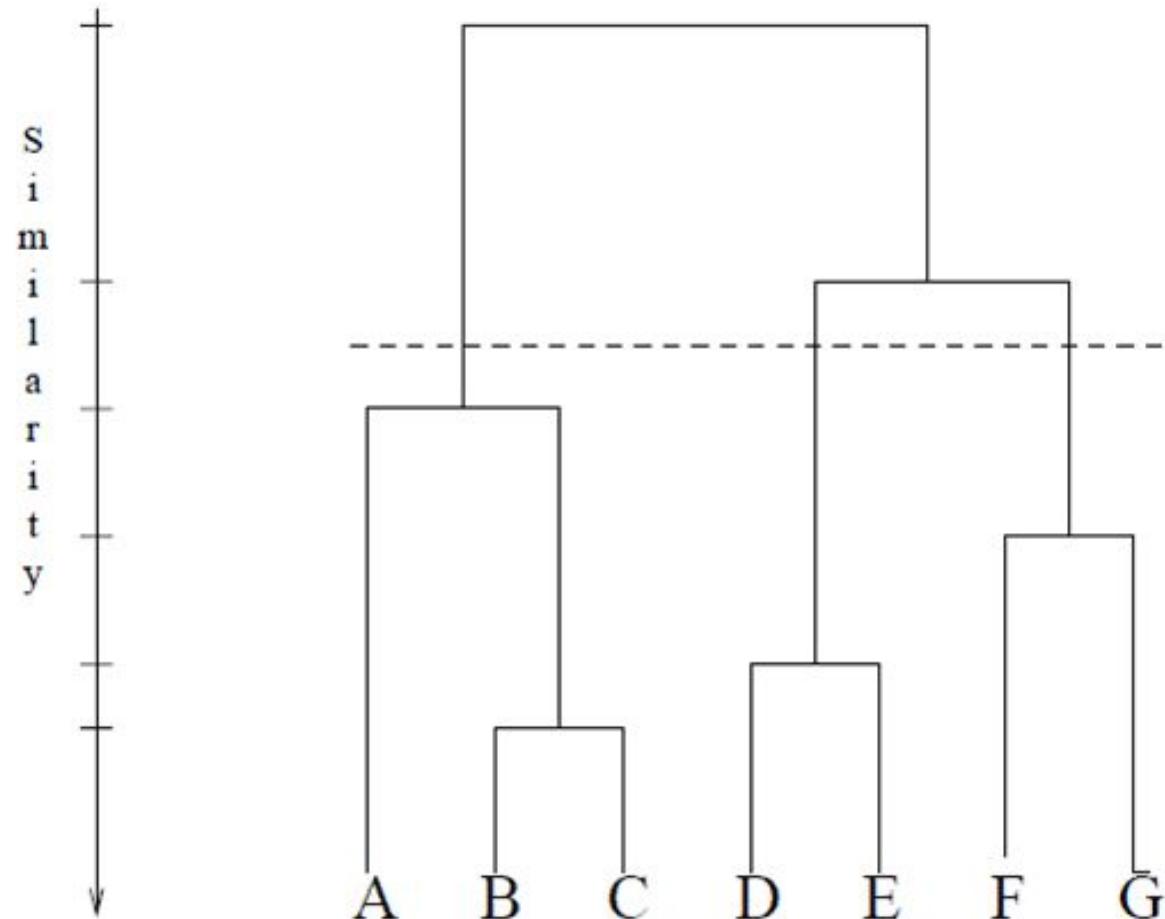


Exclusive clustering



Overlapping clustering

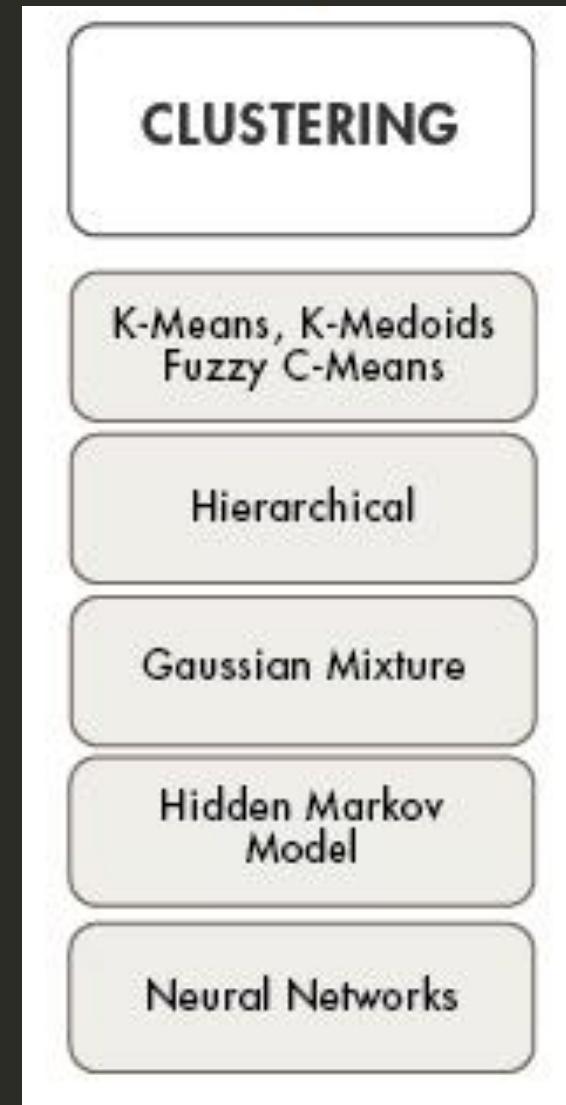




Clustering



Clustering





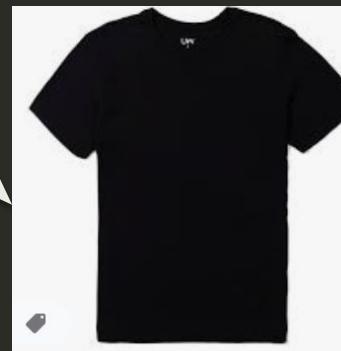
Data Science & Machine Learning



Aula 17: ML Recomendação

Recomendação

Collaborative Filtering





User ratings



10



5



1

Matrix de categorias dos produtos

Moda Masc Coleção verão Esporte fino Moda casual



Sim Sim Não Não



Sim Sim Não Não



Sim Não Não Sim

Matrix de categorias dos produtos

Moda Masc Coleção verão Esporte fino Moda casual



1 1 0 0



1 1 0 0



1 0 0 1



User ratings



10



5



1

X

Matrix de categorias dos produtos

Moda Masc Coleção verão Esporte fino Moda casual



1 1 0 0



1 1 0 0



1 0 0 1

Matrix de Categorias com Peso

Moda Masc Coleção verão Esporte fino Moda casual



10 10 0 0



5 5 0 0



1 0 0 1

Matrix de Categorias com Peso

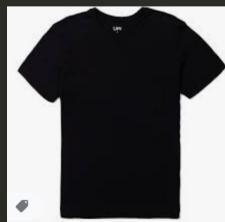
	Moda Masc	Coleção verão	Esporte fino	Moda casual
--	-----------	---------------	--------------	-------------



	10	10	0	0
--	----	----	---	---



	5	5	0	0
--	---	---	---	---



	1	0	0	1
--	---	---	---	---



User profile	16	15	1
--------------	----	----	---

User Profile

Moda Masc

Coleção verão

Esporte fino

Moda casual



16

15

0

1

User Profile

Moda Masc

Coleção verão

Esporte fino

Moda casual



16/32

15/32

0

1/32

Moda Masc

Coleção verão

Esporte fino

Moda casual



0.5

0.46

0

0.04

Moda Masc

Coleção verão

Esporte fino

Moda casual



0.5

0.46

0

0.04



Sim

Sim

Não

Não



Sim

Não

Sim

Não

Moda Masc

Coleção verão

Esporte fino

Moda casual



0.5

0.46

0

0.04



1

1

0

0



1

0

1

0

Moda Masc

Coleção verão

Esporte fino

Moda casual



0.5

0.46

0

0.04



0.5

0.46

0

0



0.5

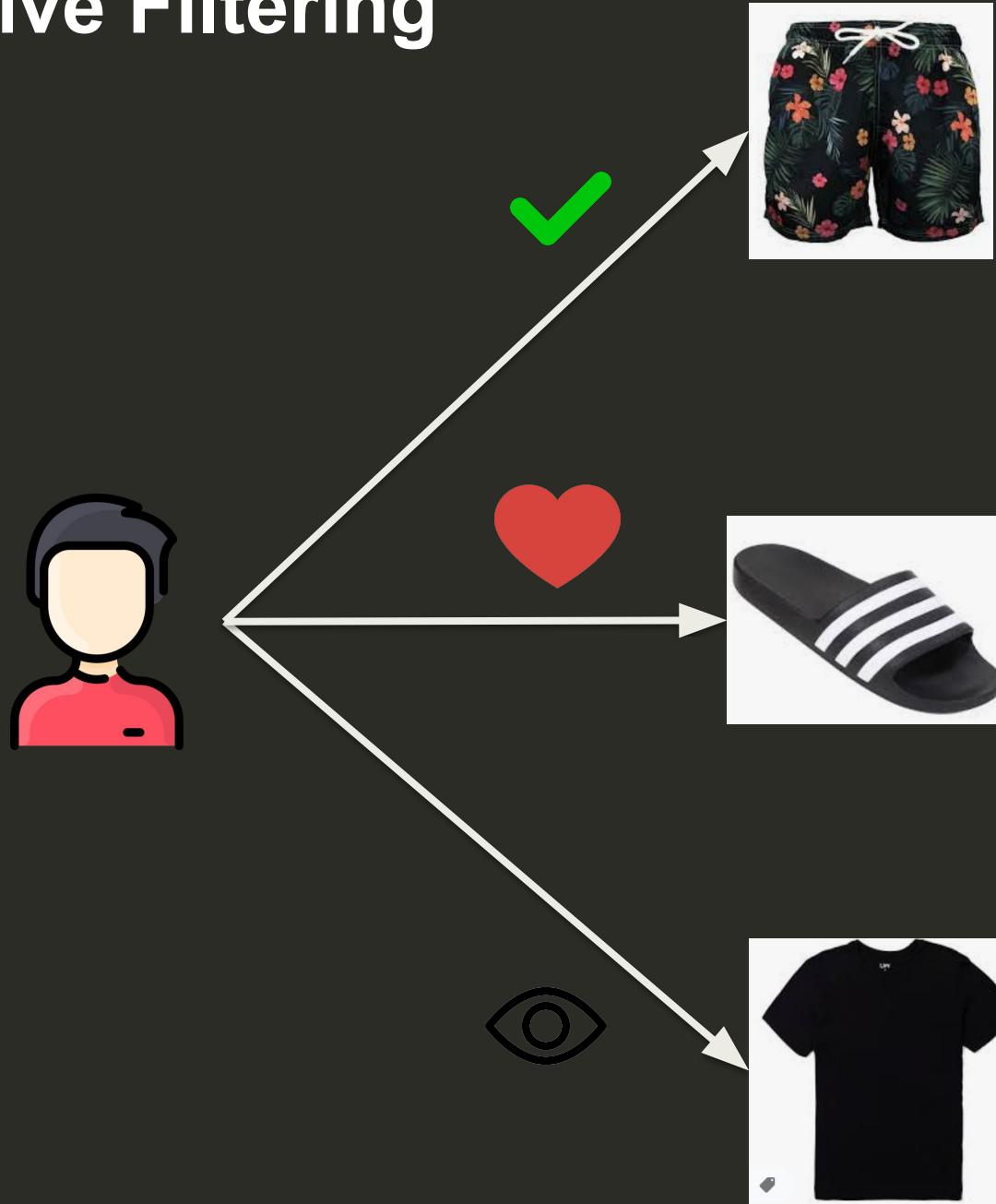
0

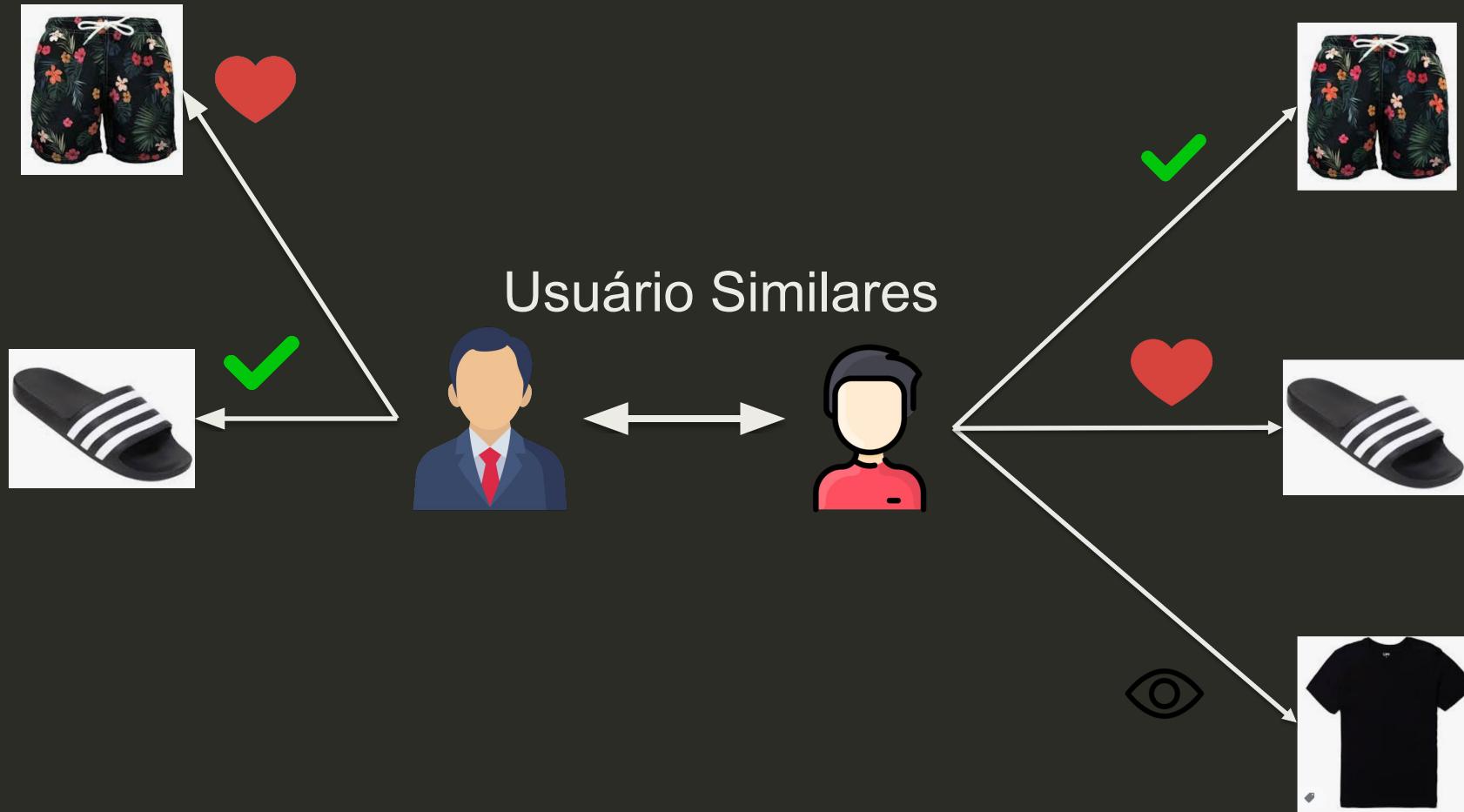
0

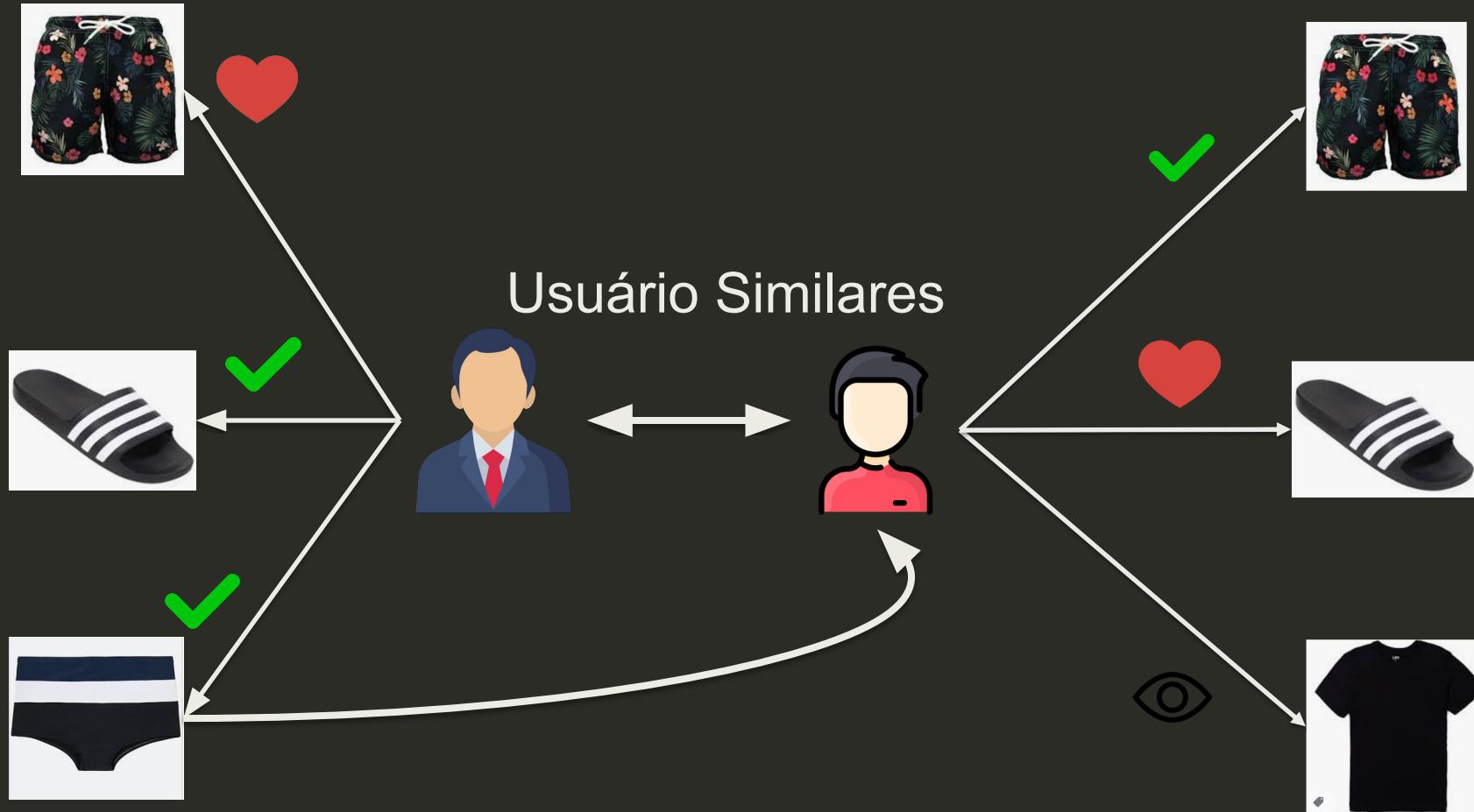
0

	Moda Masc	Coleção verão	Esporte fino	Moda casual	Total
	0.5	0.46	0	0	0.96
	0.5	0	0	0	0.5

Collaborative Filtering









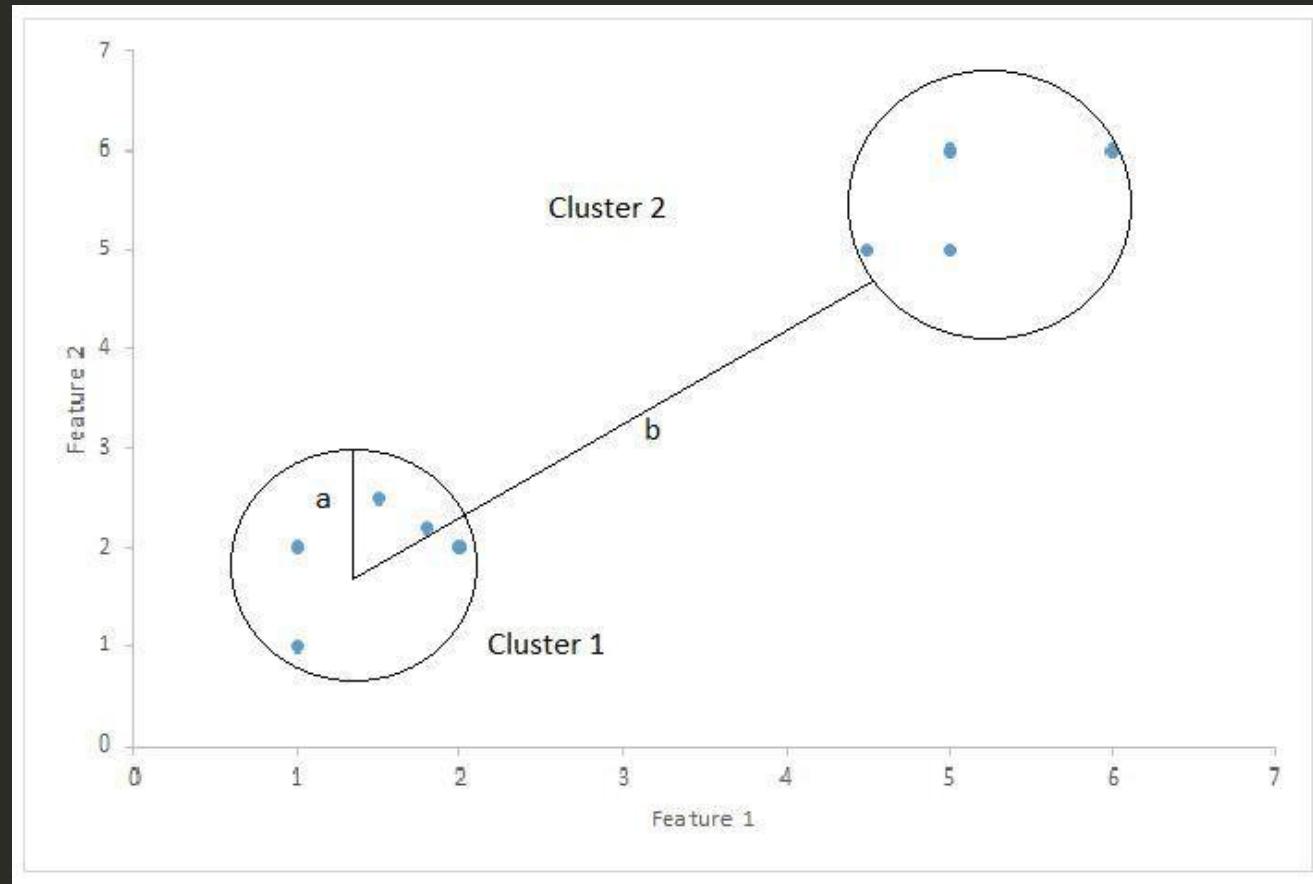
Data Science & Machine Learning



Aula 18: Métricas de avaliação de modelos Não Supervisionados

Métricas de avaliação de Clustering

Silhouette Score



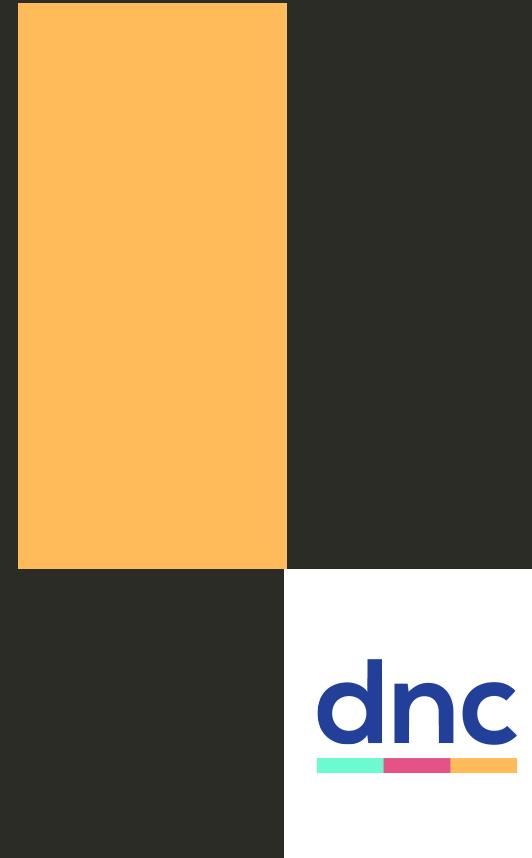
Métricas de avaliação de Recomendação

Mean Reciprocal Rank (MRR)

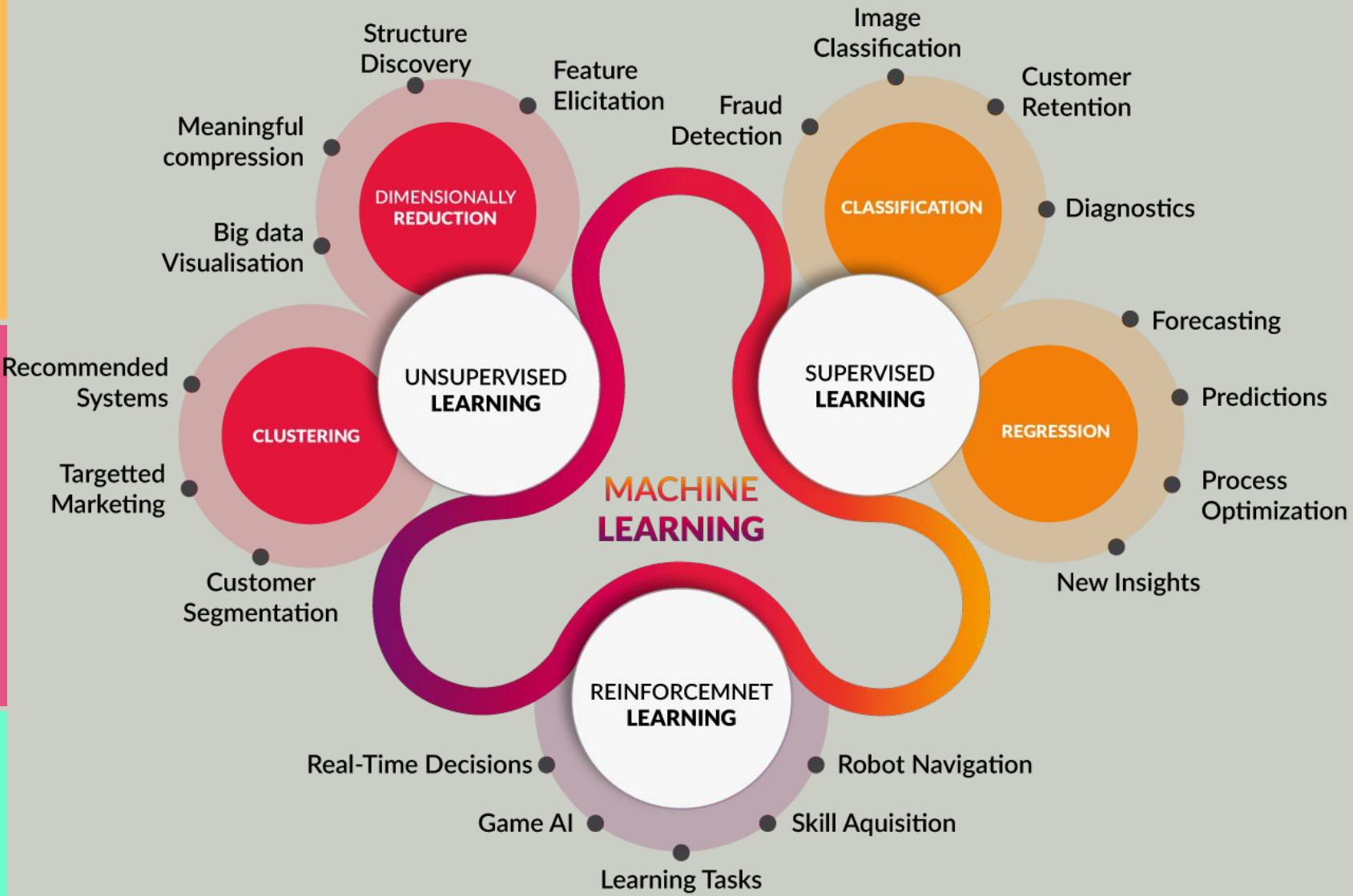
Mean Average Precision (MAP)

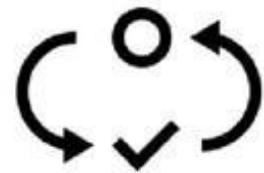


Data Science & Machine Learning



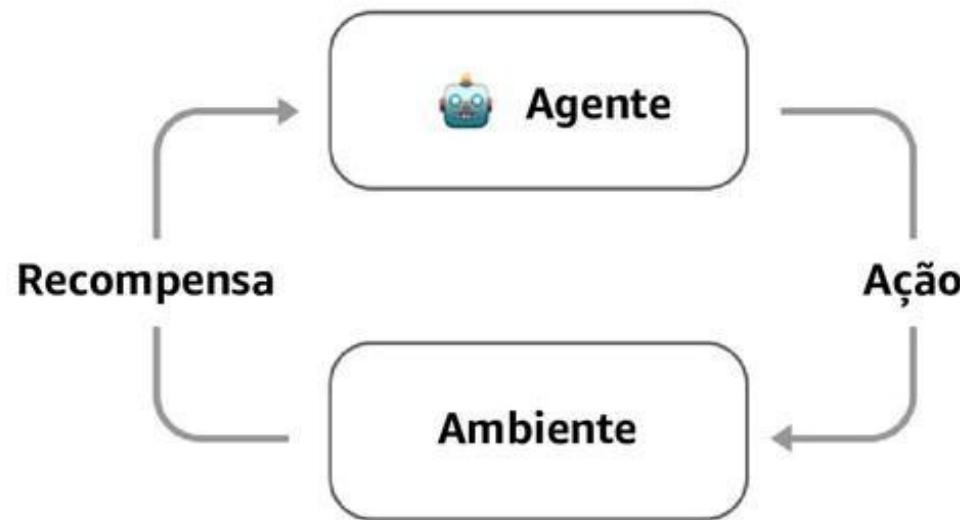
Aula 19: Machine Learning por Reforço

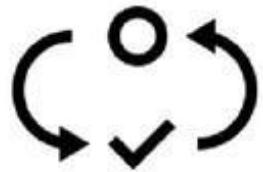




Aprendizado por Reforço

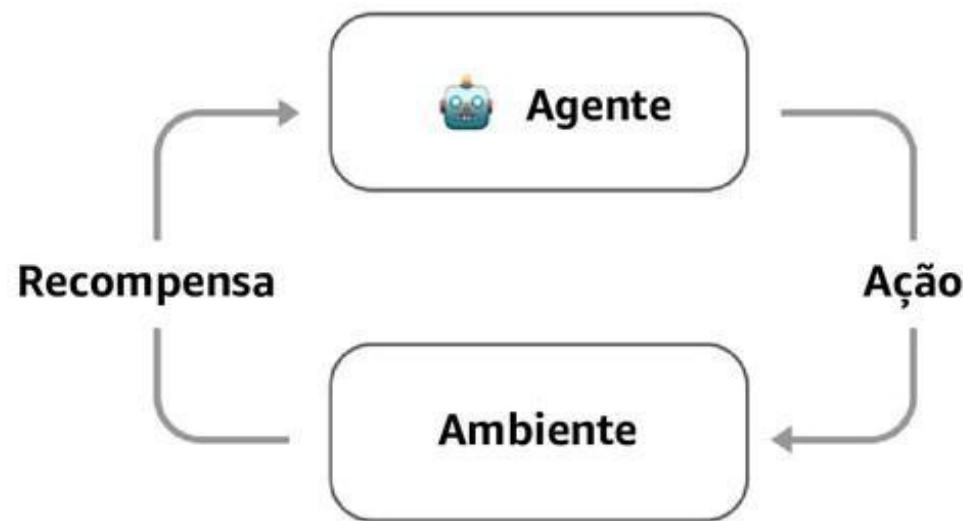
Treinar um agente para tomar certas ações dentro de um ambiente.





Aprendizado por Reforço

Treinar um agente para tomar certas ações dentro de um ambiente.

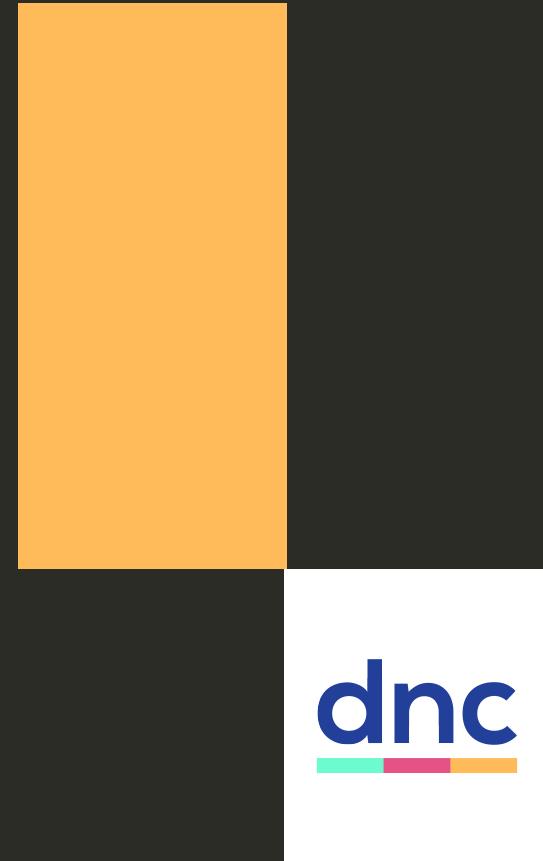


Desafios





Data Science & Machine Learning



Aula 20: AutoML

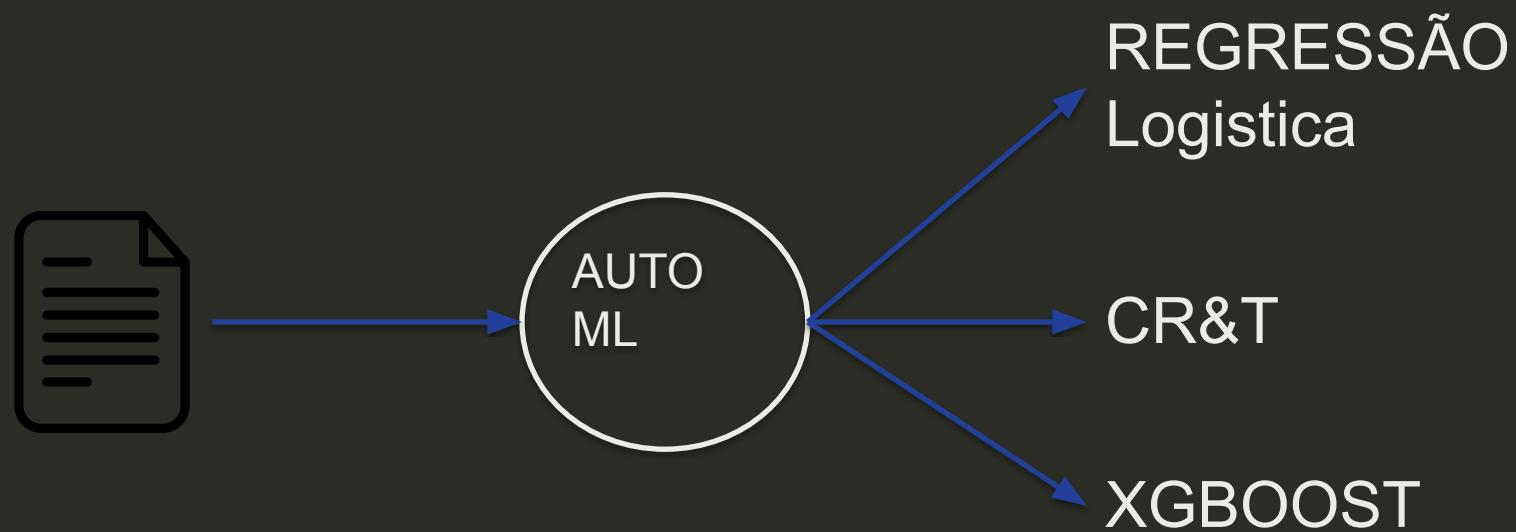
Automated ML

Autometed Filtering Enginnering

Autometed Model Selection

Automated Hyperparameter Optimization

Projeto AutoML



Automated ML

Google AUTOML

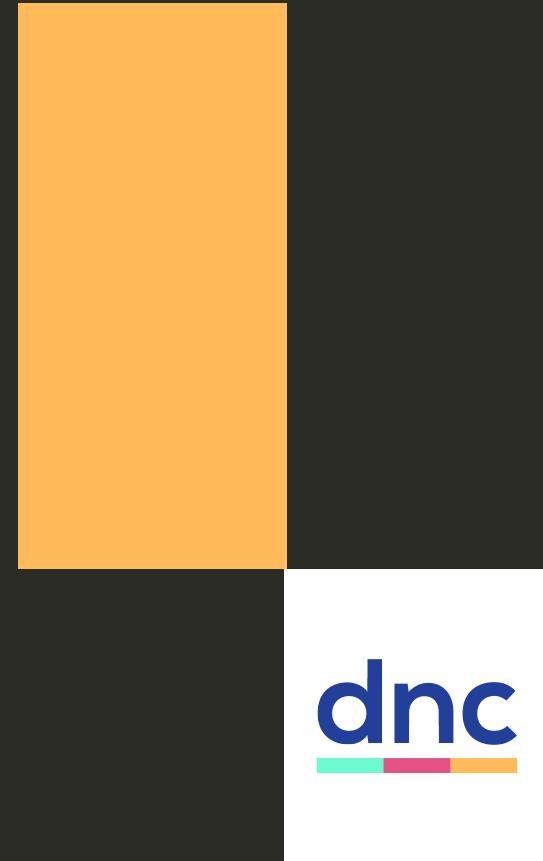
Microsoft AutoML

IBM SPSS AutoML

H2O AI AutoML



Data Science & Machine Learning



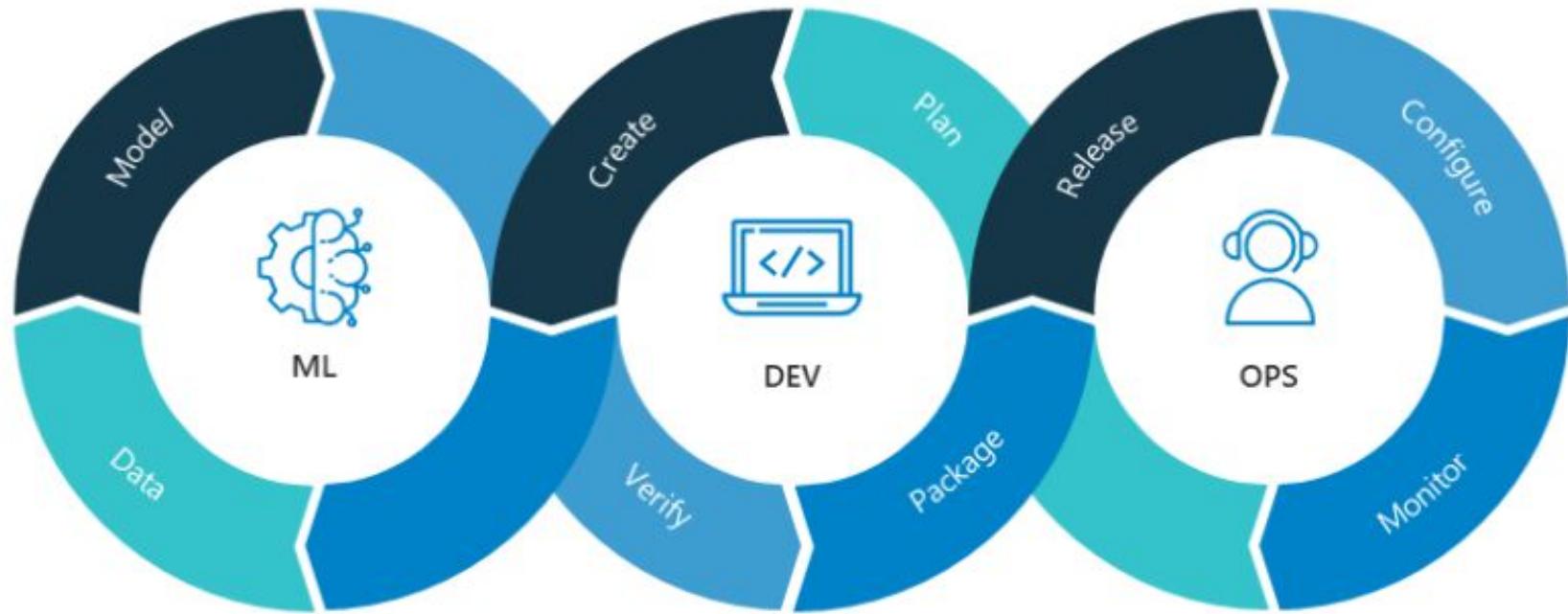
Aula 21: ML Ops

DEV Ops

DevOps é uma prática popular no desenvolvimento e operação de sistemas de software em grande escala.

CI – Continuos Integration

CD – Continuos Delivery



ML Ops

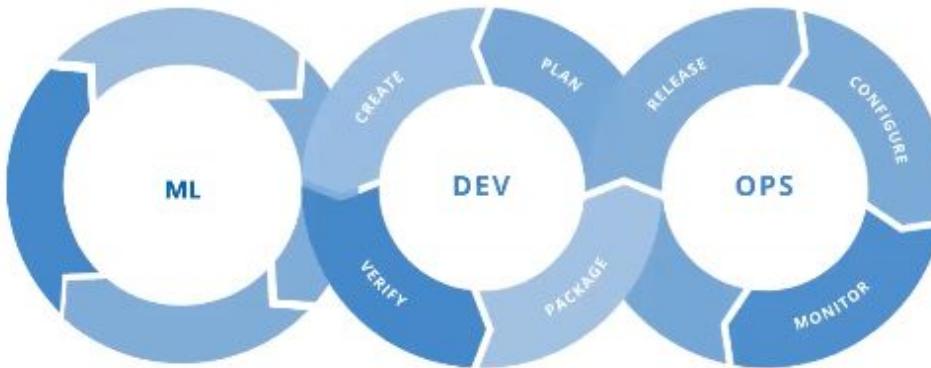
CI – Continuos Integration

CD – Continuos Delivery

CT – Continuos Trainning



MLOps = ML + DEV + OPS



Experiment

Data Acquisition
Business Understanding
Initial Modeling

Develop

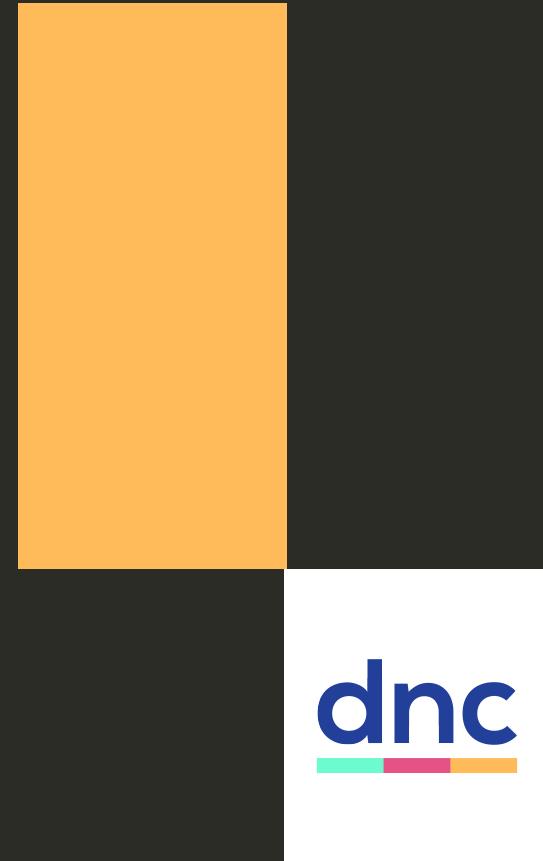
Modeling + Testing
Continuous Integration
Continuous Deployment

Operate

Continuous Delivery
Data Feedback Loop
System + Model Monitoring



Data Science & Machine Learning



Aula 22: Vieses de IA (Bias)

Víeses





Dr. Alex Hanna / Kate Silver

@alexhanna



Apparently Twitter uses a facial analysis algorithm to center image previews. This quick little robustness test shows how it centers the white guy in both images.



nota @NotAFile

Replying to @taintedavenue @chochosmx and 2 others

There you go

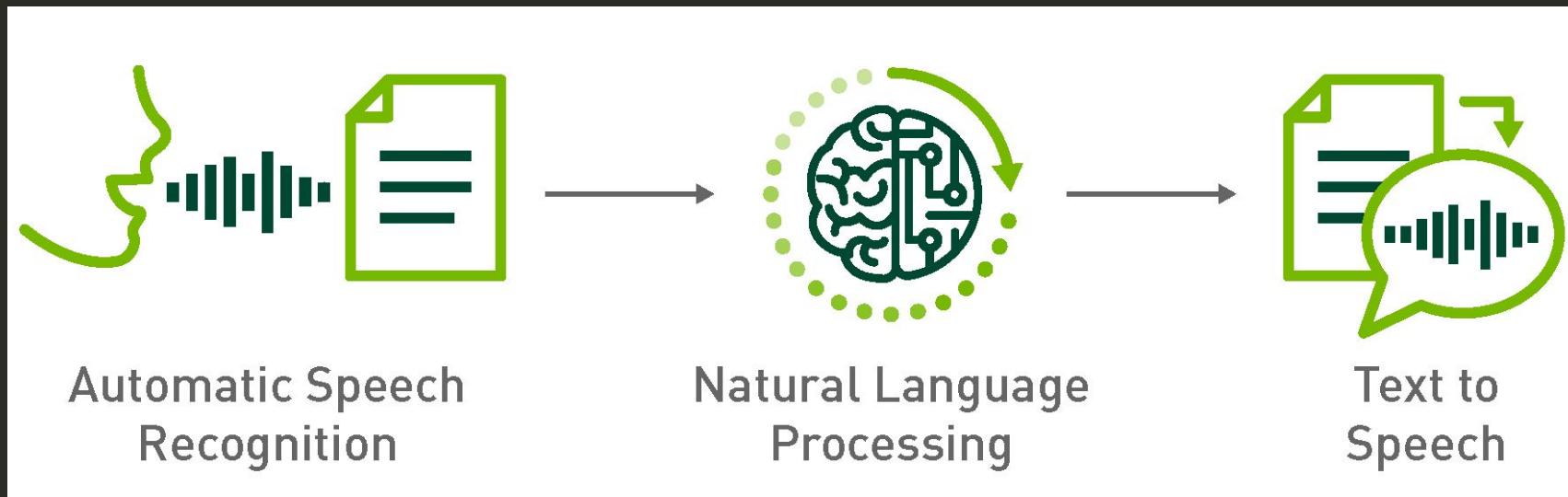


9:12 PM · Sep 19, 2020



♡ 4.6K

Viéses



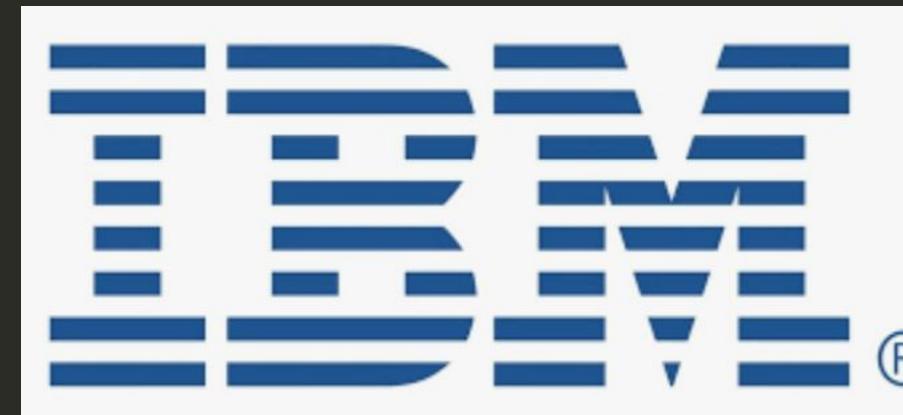
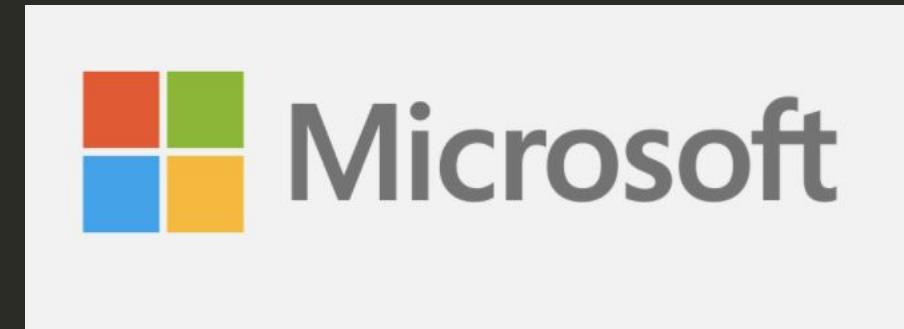
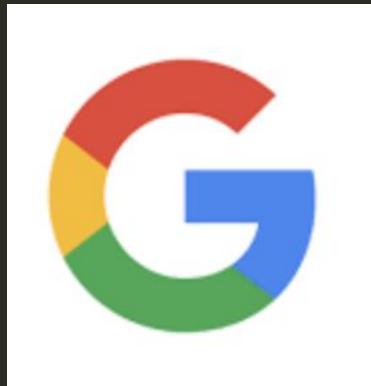
Viéses

<https://aif360.mybluemix.net/>

AI Fairness 360

This extensible open source toolkit can help you examine, report, and mitigate discrimination and bias in machine learning models throughout the AI application lifecycle. We invite you to use and improve it.

Ética de IA





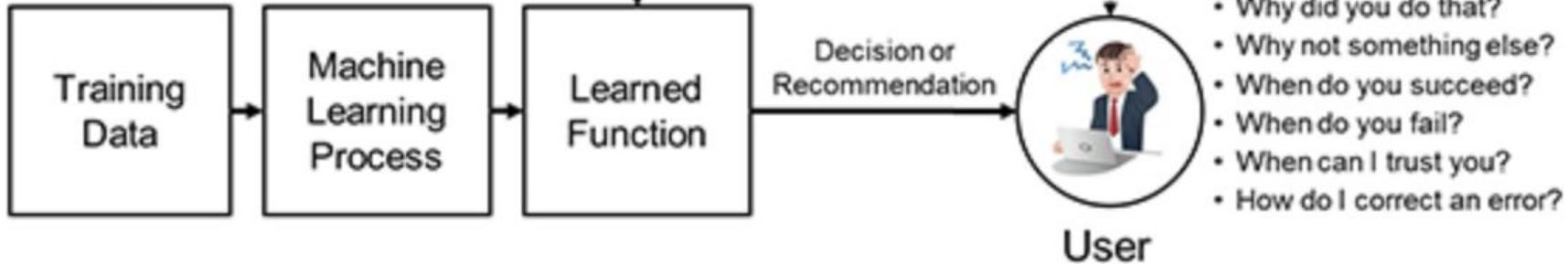
Data Science & Machine Learning



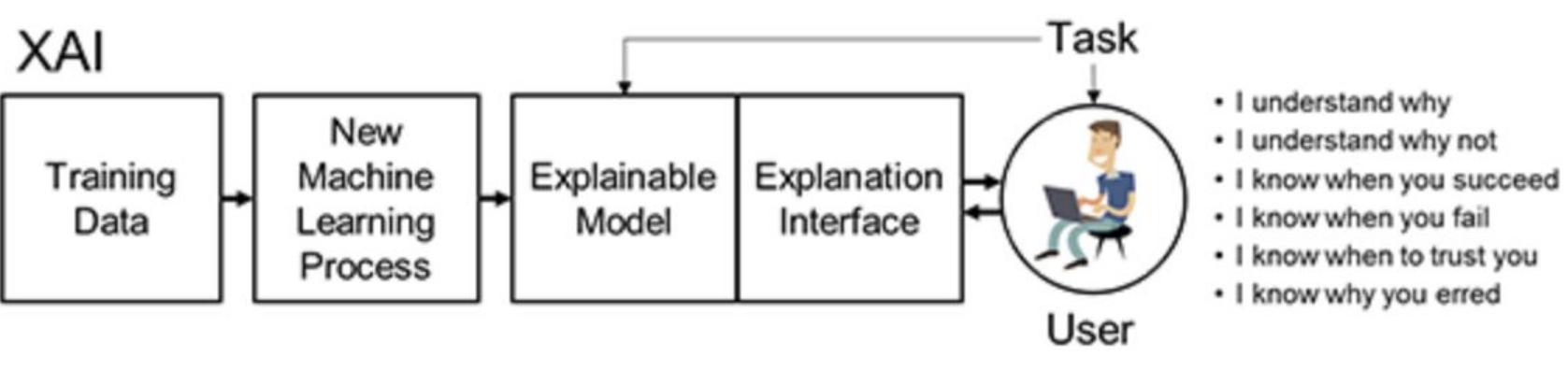
Aula 23: Desafios da IA (Explicabilidade)

Explicabilidade

Today



Explicabilidade (XAI)



Explicabilidade

<https://aix360.mybluemix.net/>

AI Explainability 360

This extensible open source toolkit can help you comprehend how machine learning models predict labels by various means throughout the AI application lifecycle. We invite you to use it and improve it.

Explicabilidade

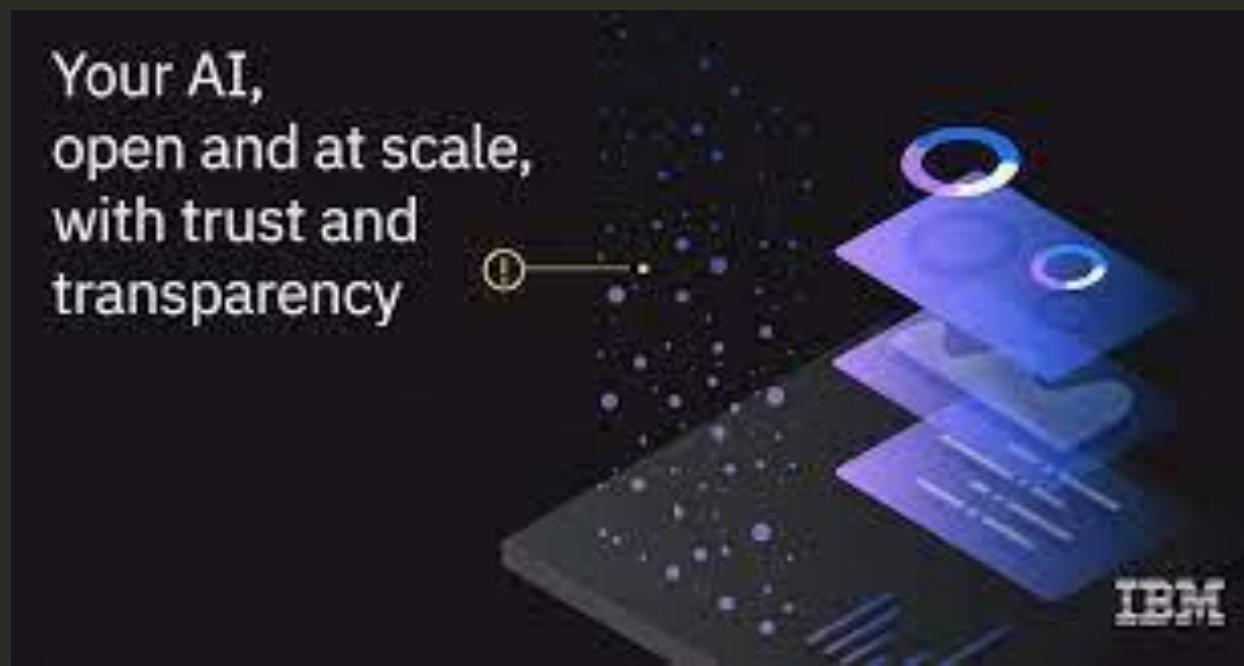
<https://lime-ml.readthedocs.io/en/latest/>

Local Interpretable Model-Agnostic Explanations (lime)

In this page, you can find the Python API reference for the lime package (local interpretable model-agnostic explanations). For tutorials and more information, visit [the github page](#).

Explicabilidade

Watson Open Scale



Explicabilidade (XAI)

35% a 50%

Esforço reduzido de monitoramento de modelo

Entre 3x e 8x

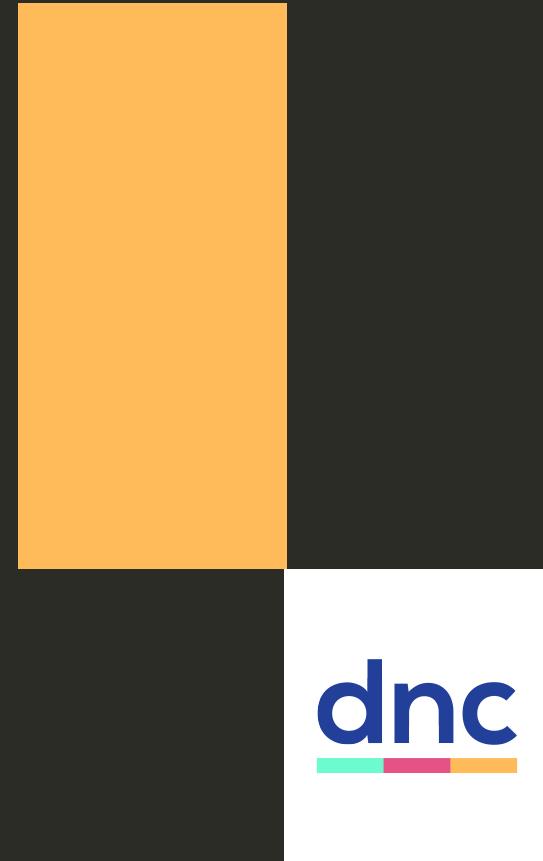
Aumento dos modelos em produção

15% a 30%

Maior precisão dos modelos



Data Science & Machine Learning



Aula 24: IA no Dia a Dia

NLP



Voz



Imagen

Classificação de Imagem - OCR



Imagen

Reconhecimento de Face. Facebook – Desbloqueio do celular

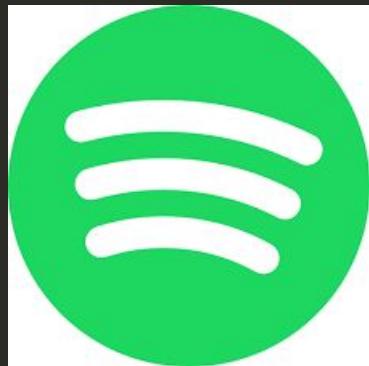


Imagen

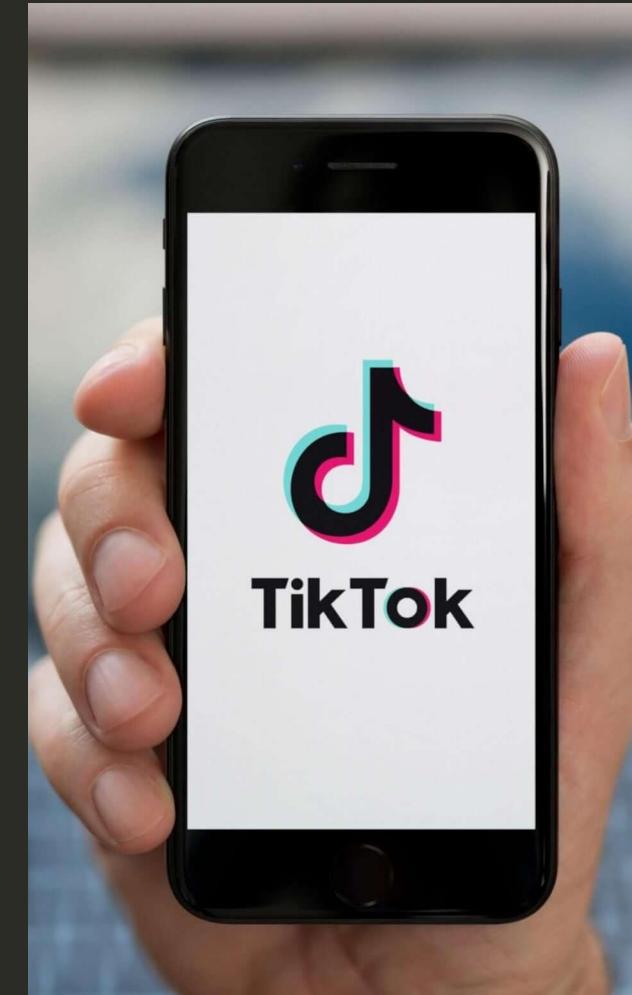
Reconhecimento de Face. Facebook – Desbloqueio do celular



RECOMENDAÇÃO



NETFLIX



ANÁLISE DE CRÉDITO

Análise de crédito do Nubank

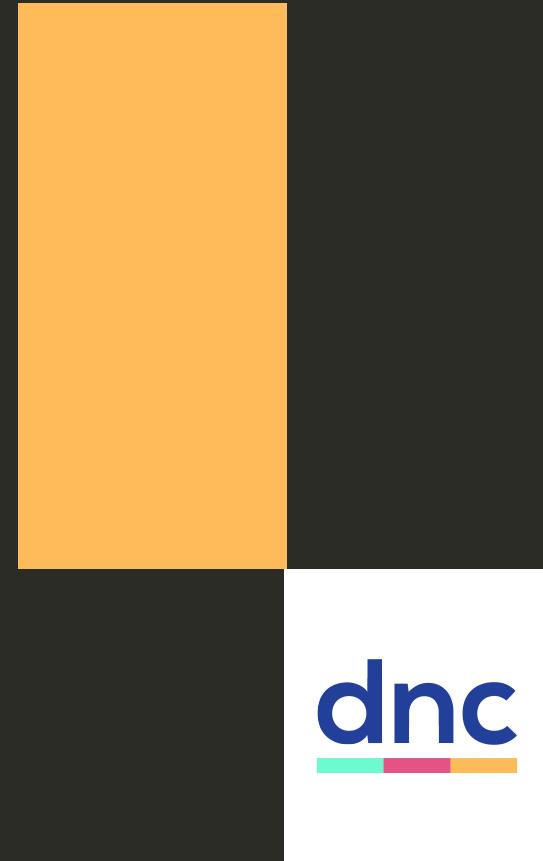


ABERTURA DE CONTAS DIGITAL





Data Science & Machine Learning



Aula 30: Soluções de IA

AGRICULTURA

BANCOS E MERCADO FINANCEIRO

EDUCAÇÃO

AVIAÇÃO

INDÚSTRIA ALIMENTÍCIA/BEBIDAS

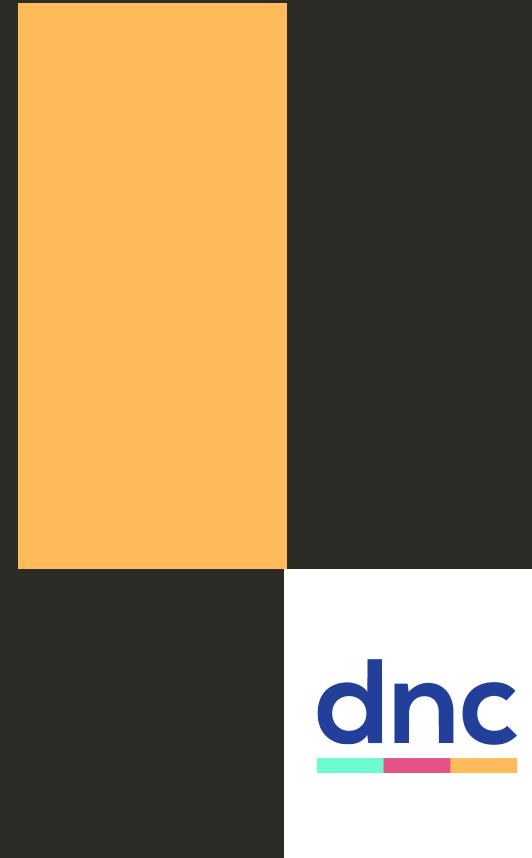
TRANSPORTE/SEGURANÇA

SAÚDE E MEDICINA

VAREJO



Data Science & Machine Learning



Aula 25: Introdução a Deep Learning



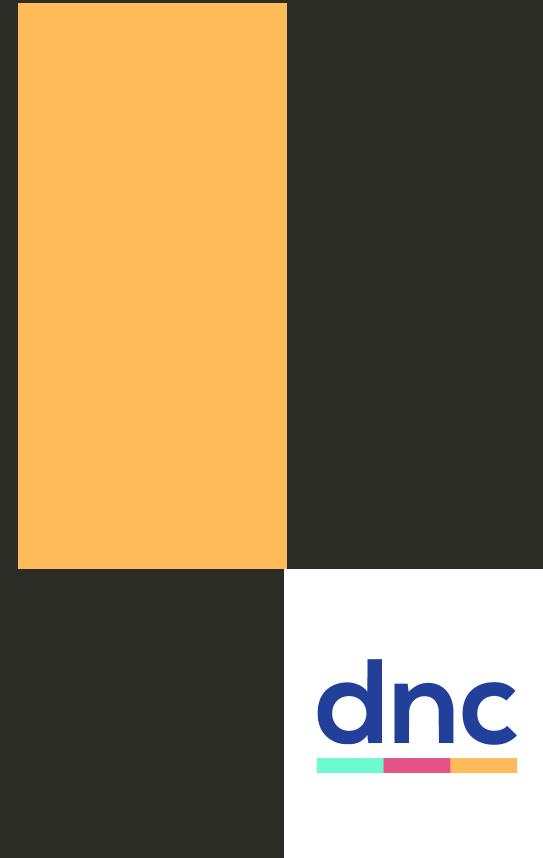
Data Science & Machine Learning



Aula 26: Redes Convolucionais



Data Science & Machine Learning



Aula 27: Redes Recorrentes



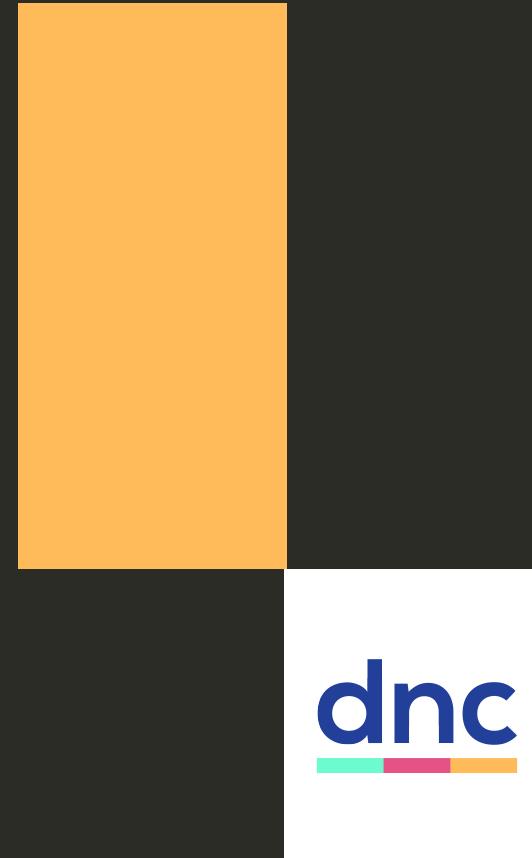
Data Science & Machine Learning



Aula 28: Tensor Flow, Keras e Pytorch



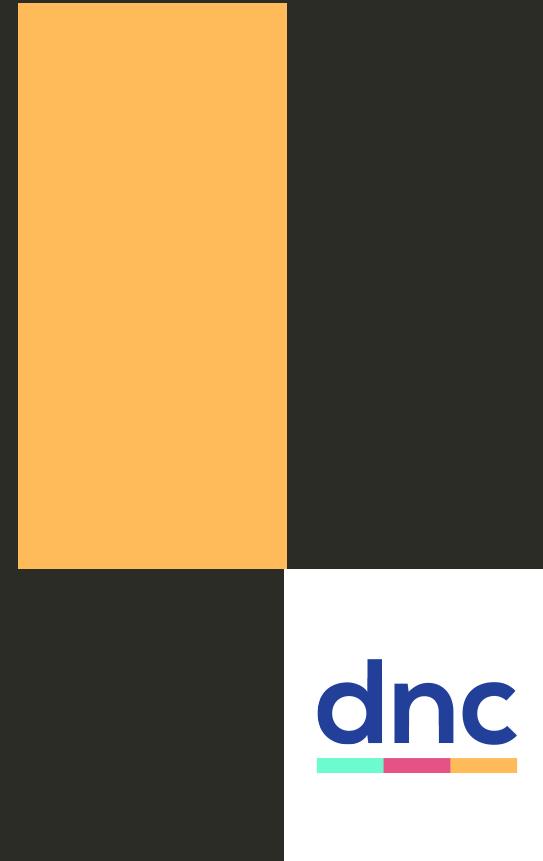
Data Science & Machine Learning



Aula 29: One Shot Learning



Data Science & Machine Learning



Aula 31: Resumo

2	Introdução a Inteligência Artificial + História de IA
3	Subáreas da Inteligência Artificial
4	Chatbots e Agentes Virtuais
5	NLP
6	Reconhecimento de Voz
7	Visão Computacional
8	Serviços e APIs de Inteligência Artificial
Não feito	Quiz - Introdução a Machine Learning
9	Inteligência Artificial para Ciência de Dados
10	Machine Learning Supervisionado
11	Machine Learning Supervisionado: Regressão
12	Machine Learning Supervisionado: Classificação
13	Métodos de Avaliação de Machine Learning supervisionado
14	Underfitting e Overfitting + Como evitar
Não feito	Quiz - Introdução a Machine Learning
15	Machine Learning Não Supervisionado
16	Machine Learning Não-Supervisionado: Clustering
17	Machine Learning Não-Supervisionado: Recomendação
18	Métricas de Avaliação de Machine Learning Não-Supervisionado
19	Machine Learning Não-Supervisionado: Reforço
Não feito	Quiz - Introdução a Machine Learning
20	AutoML
21	ML Ops
22	Vieses de Inteligência Artificial (Bias)
23	Desafios de Inteligência Artificial (Explicabilidade)
24	Inteligência Artificial no Dia-a-Dia



Data Science & Machine Learning



Aula 32: Como estudar I.A.

COMO LIDAR COM TERMOS EM INGLÊS

COMO LIDAR COM TERMOS QUE NÃO CONHEÇO

COMO SE MANTER ATUALIZADO

You can never understand everything. But, you should push yourself to understand the system.

- Ryan Dahl (Creator of Node JS)