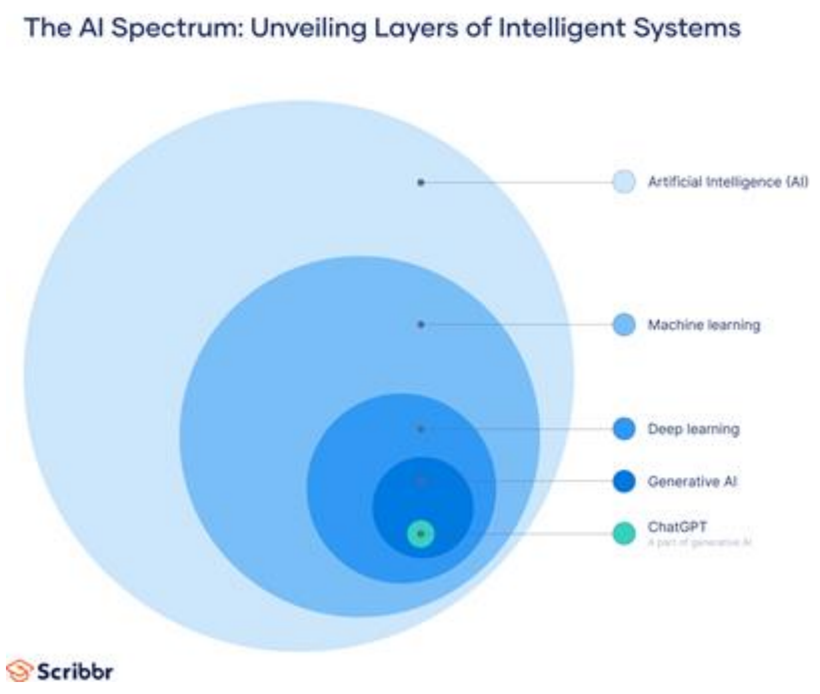
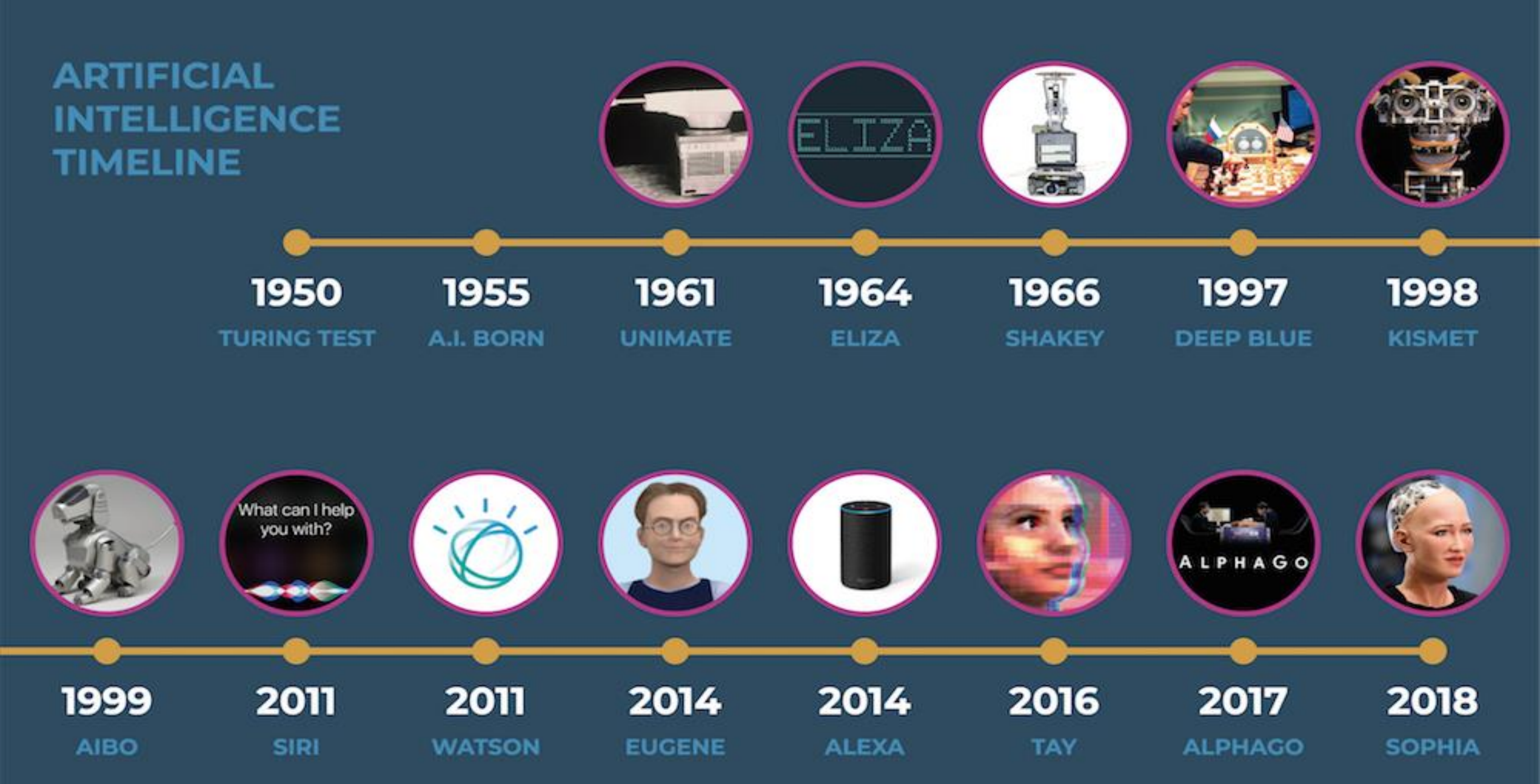
Aprendizado Supervisionado

# INTRODUÇÃO AO APRENDIZADO SUPERVISIONADO

Marcelo Mine




Fonte: what is deep learning? | A beginner's guide




CRÉDITO: OBI4WAN

# / THIS PERSON DOES NOT EXIST

https://thispersondoesnotexist.com/



/ Watson and the jeopardy

/ A Eurovision song created by AI


/ Kungs vs Cookin' on 3 Burners - This Girl

# / VIESES

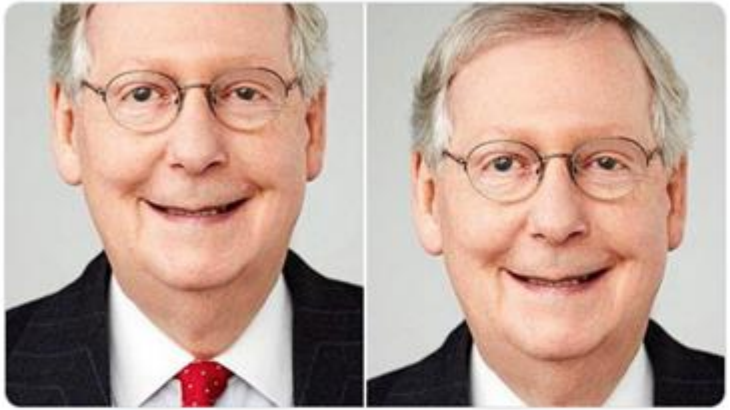
https://twitter.com/nke_ise/status/897756900753891328?s=20

# / VIESES

Tony "Abolish ICE" Arcieri 🦀 🌹
@bascule

Trying a horrible experiment...

Which will the Twitter algorithm pick: Mitch McConnell or Barack Obama?



7:05 PM · Sep 19, 2020

## Twitter apologises for 'racist' image-cropping algorithm

**Users highlight examples of feature automatically focusing on white faces over black ones**

▲ Twitter users began to spot flaws in the feature over the weekend. Photograph: Glenn Chapman/AFP/Getty Images

FONTE: THE GUARDIAN

# / CASOS DE USO: THE BIG BOOK OF DATA SCIENCE USE CASES

## Contents

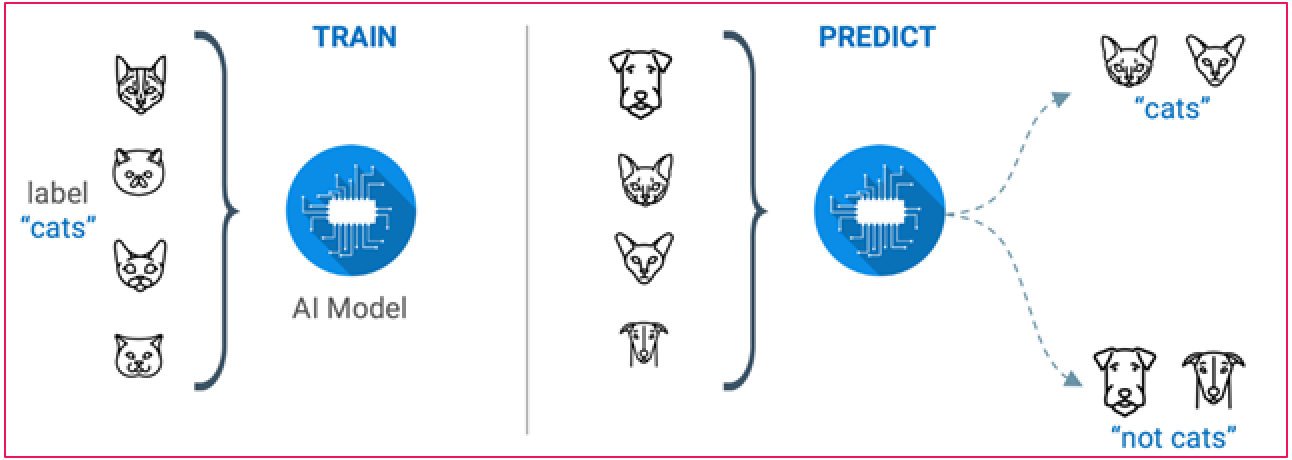
# / CURSO

# CONTEÚDO

# / SOBRE O CURSO

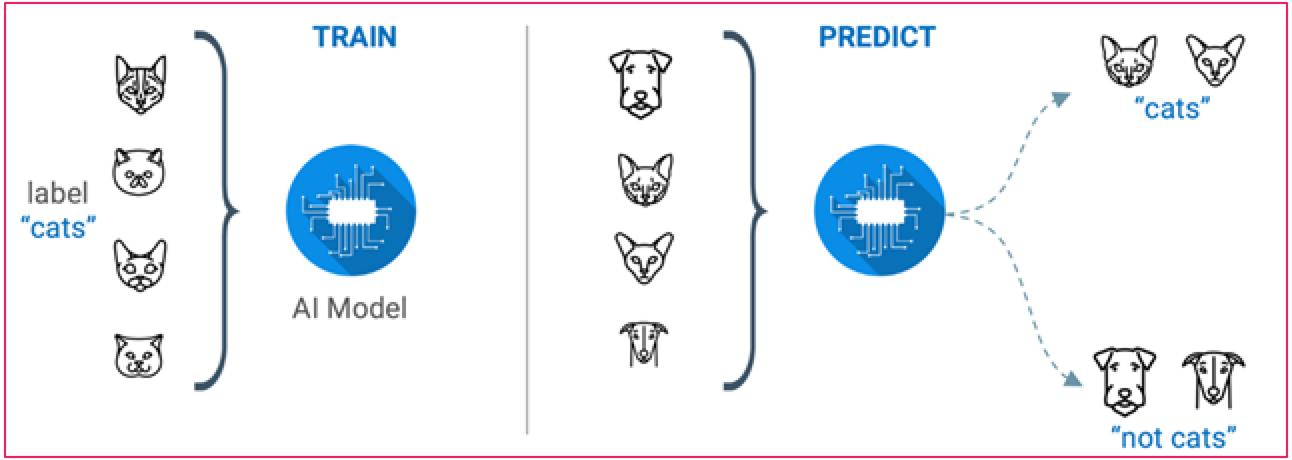
- Hands-On
- Python 3
- Bibliotecas / Pacotes / Módulos

# APRENDIZADO SUPERVISIONADO



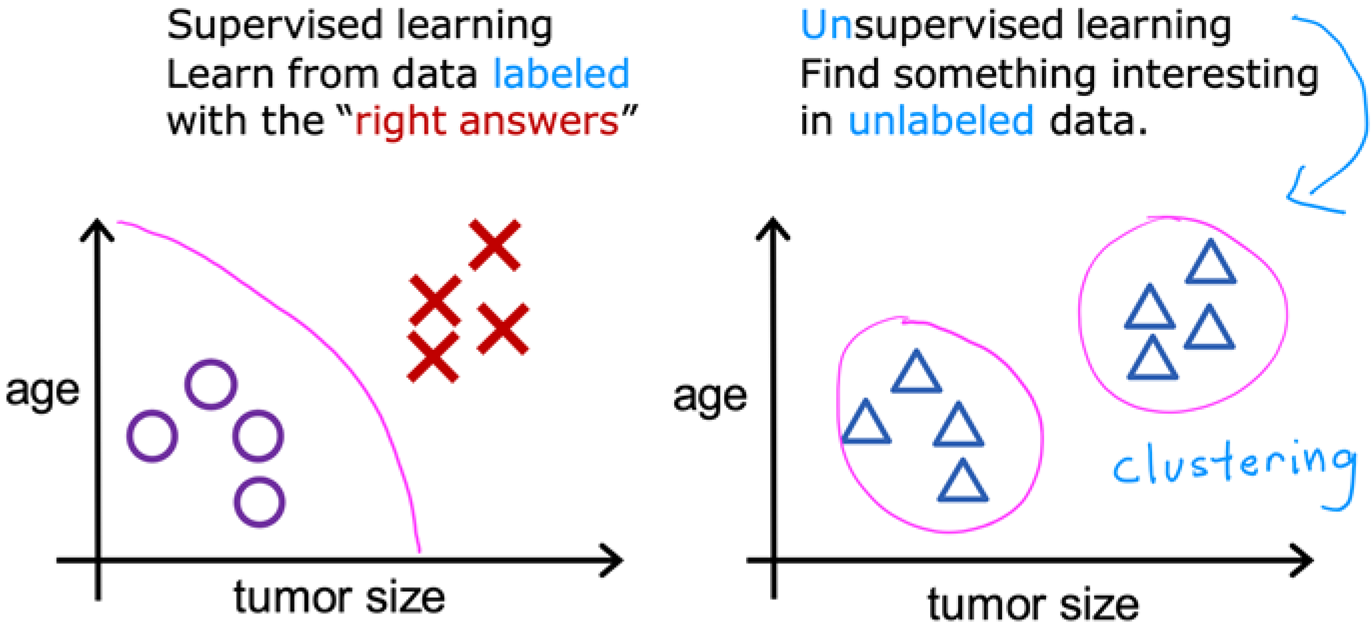

FONTE IMAGEM: ABEYON

- Técnica de aprendizado de máquina em que um algoritmo aprende a partir de dados rotulados
- A máquina aprende a fazer previsões ou decisões com base nos recursos de entrada e nos rótulos de saída correspondentes
- Ele requer um conjunto de dados de treinamento com exemplos rotulados para aprender
- É usado quando queremos prever um determinado resultado a partir de uma determinada entrada e temos exemplos de pares entrada/saída. Nosso objetivo é realizar previsões precisas para dados novos e nunca antes vistos.




FONTE IMAGEM: ABEYON

- A aprendizagem supervisionada muitas vezes requer esforço humano para construir o conjunto de treinamento, mas depois automatiza e muitas vezes acelera uma tarefa que de outra forma seria trabalhosa ou inviável.
- Diz respeito a tarefas como **classificação** e **regressão.**




Fonte: [Machine Learning Specialization]

# / CLASSIFICAÇÃO

Classe ser do tipo discreto (valores inteiros ou categorias)

| User ID | Gender | Age | Salary | Purchased |
|---|---|---|---|---|
| 15624510 | Male | 19 | 19000 | 0 |
| 15810944 | Male | 35 | 20000 | 1 |
| 15668575 | Female | 26 | 43000 | 0 |
| 15603246 | Female | 27 | 57000 | 0 |
| 15804002 | Male | 19 | 76000 | 1 |
| 15728773 | Male | 27 | 58000 | 1 |
| 15598044 | Female | 27 | 84000 | 0 |
| 15694829 | Female | 32 | 150000 | 1 |
| 15600575 | Male | 25 | 33000 | 1 |
| 15727311 | Female | 35 | 65000 | 0 |
| 15570769 | Female | 26 | 80000 | 1 |
| 15606274 | Female | 26 | 52000 | 0 |
| 15746139 | Male | 20 | 86000 | 1 |
| 15704987 | Male | 32 | 18000 | 0 |
| 15628972 | Male | 18 | 82000 | 0 |
| 15697686 | Male | 29 | 80000 | 0 |
| 15733883 | Male | 47 | 25000 | 1 |

Fonte: Supervised machine learning

# / REGRESSÃO

Classe possui valores contínuos (valores no intervalo dos números reais)

| Temperature | Pressure | Relative Humidity | Wind Direction | Wind Speed |
|---|---|---|---|---|
| 10.69261758 | 986.882019 | 54.19337313 | 195.7150879 | 3.278597116 |
| 13.59184184 | 987.8729248 | 48.0648859 | 189.2951202 | 2.909167767 |
| 17.70494885 | 988.1119385 | 39.11965597 | 192.9273834 | 2.973036289 |
| 20.95430404 | 987.8500366 | 30.66273218 | 202.0752869 | 2.965289593 |
| 22.9278274 | 987.2833862 | 26.06723423 | 210.6589203 | 2.798230886 |
| 24.04233986 | 986.2907104 | 23.46918024 | 221.1188507 | 2.627005816 |
| 24.41475295 | 985.2338867 | 22.25082295 | 233.7911987 | 2.448749781 |
| 23.93361956 | 984.8914795 | 22.35178837 | 244.3504333 | 2.454271793 |
| 22.68800023 | 984.8461304 | 23.7538641 | 253.0864716 | 2.418341875 |
| 20.56425726 | 984.8380737 | 27.07867944 | 264.5071106 | 2.318677425 |
| 17.76400389 | 985.4262085 | 33.54900114 | 280.7827454 | 2.343950987 |
| 11.25680746 | 988.9386597 | 53.74139903 | 68.15406036 | 1.650191426 |
| 14.37810685 | 989.6819458 | 40.70884681 | 72.62069702 | 1.553469896 |
| 18.45114201 | 990.2960205 | 30.85038484 | 71.70604706 | 1.005017161 |
| 22.54895853 | 989.9562988 | 22.81738811 | 44.66042709 | 0.264133632 |
| 24.23155922 | 988.796875 | 19.74790765 | 318.3214111 | 0.329656571 |

Fonte: Supervised machine learning

# / CLASSIFICAÇÃO OU REGRESSÃO?

| | age | sex | chest pain type | resting bp s | cholesterol | fasting blood sugar | resting ecg | max heart rate | exercise angina | oldpeak | ST slope | target |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 40 | 1 | 2 | 140 | 289 | 0 | 0 | 172 | 0 | 0.0 | 1 | 0 |
| 1 | 49 | 0 | 3 | 160 | 180 | 0 | 0 | 156 | 0 | 1.0 | 2 | 1 |
| 2 | 37 | 1 | 2 | 130 | 283 | 0 | 1 | 98 | 0 | 0.0 | 1 | 0 |
| 3 | 48 | 0 | 4 | 138 | 214 | 0 | 0 | 108 | 1 | 1.5 | 2 | 1 |
| 4 | 54 | 1 | 3 | 150 | 195 | 0 | 0 | 122 | 0 | 0.0 | 1 | 0 |
| 5 | 39 | 1 | 3 | 120 | 339 | 0 | 0 | 170 | 0 | 0.0 | 1 | 0 |
| 6 | 45 | 0 | 2 | 130 | 237 | 0 | 0 | 170 | 0 | 0.0 | 1 | 0 |
| 7 | 54 | 1 | 2 | 110 | 208 | 0 | 0 | 142 | 0 | 0.0 | 1 | 0 |
| 8 | 37 | 1 | 4 | 140 | 207 | 0 | 0 | 130 | 1 | 1.5 | 2 | 1 |
| 9 | 48 | 0 | 2 | 120 | 284 | 0 | 0 | 120 | 0 | 0.0 | 1 | 0 |

Fonte: Kaggle

# CLASSIFICAÇÃO OU REGRESSÃO?

| | age | sex | chest pain type | resting bp s | cholesterol | fasting blood sugar | resting ecg | max heart rate | exercise angina | oldpeak | ST slope | target |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 40 | 1 | 2 | 140 | 289 | 0 | 0 | 172 | 0 | 0.0 | 1 | 0 |
| 1 | 49 | 0 | 3 | 160 | 180 | 0 | 0 | 156 | 0 | 1.0 | 2 | 1 |
| 2 | 37 | 1 | 2 | 130 | 283 | 0 | 1 | 98 | 0 | 0.0 | 1 | 0 |
| 3 | 48 | 0 | 4 | 138 | 214 | 0 | 0 | 108 | 1 | 1.5 | 2 | 1 |
| 4 | 54 | 1 | 3 | 150 | 195 | 0 | 0 | 122 | 0 | 0.0 | 1 | 0 |
| 5 | 39 | 1 | 3 | 120 | 339 | 0 | 0 | 170 | 0 | 0.0 | 1 | 0 |
| 6 | 45 | 0 | 2 | 130 | 237 | 0 | 0 | 170 | 0 | 0.0 | 1 | 0 |
| 7 | 54 | 1 | 2 | 110 | 208 | 0 | 0 | 142 | 0 | 0.0 | 1 | 0 |
| 8 | 37 | 1 | 4 | 140 | 207 | 0 | 0 | 130 | 1 | 1.5 | 2 | 1 |
| 9 | 48 | 0 | 2 | 120 | 284 | 0 | 0 | 120 | 0 | 0.0 | 1 | 0 |

Fonte: Kaggle

# / CLASSIFICAÇÃO OU REGRESSÃO?

| | YearsExperience | Salary |
|---|---|---|
| 0 | 1.2 | 39344.0 |
| 1 | 1.4 | 46206.0 |
| 2 | 1.6 | 37732.0 |
| 3 | 2.1 | 43526.0 |
| 4 | 2.3 | 39892.0 |
| 5 | 3.0 | 56643.0 |
| 6 | 3.1 | 60151.0 |
| 7 | 3.3 | 54446.0 |
| 8 | 3.3 | 64446.0 |
| 9 | 3.8 | 57190.0 |

Fonte: Kaggle

# / CLASSIFICAÇÃO OU **REGRESSÃO**?

| | YearsExperience | Salary |
|---|---|---|
| 0 | 1.2 | 39344.0 |
| 1 | 1.4 | 46206.0 |
| 2 | 1.6 | 37732.0 |
| 3 | 2.1 | 43526.0 |
| 4 | 2.3 | 39892.0 |
| 5 | 3.0 | 56643.0 |
| 6 | 3.1 | 60151.0 |
| 7 | 3.3 | 54446.0 |
| 8 | 3.3 | 64446.0 |
| 9 | 3.8 | 57190.0 |

Fonte: Kaggle

# / CLASSIFICAÇÃO OU REGRESSÃO?

| | comprim_sepala | largura_sepala | comprim_petala | largura_petala | classe |
|---|---|---|---|---|---|
| 0 | 5.1 | 3.5 | 1.4 | 0.2 | Iris-setosa |
| 1 | 4.9 | 3.0 | 1.4 | 0.2 | Iris-setosa |
| 2 | 4.7 | 3.2 | 1.3 | 0.2 | Iris-setosa |
| 3 | 4.6 | 3.1 | 1.5 | 0.2 | Iris-setosa |
| 4 | 5.0 | 3.6 | 1.4 | 0.2 | Iris-setosa |
| 5 | 5.4 | 3.9 | 1.7 | 0.4 | Iris-setosa |
| 50 | 7.0 | 3.2 | 4.7 | 1.4 | Iris-versicolor |
| 51 | 6.4 | 3.2 | 4.5 | 1.5 | Iris-versicolor |
| 52 | 6.9 | 3.1 | 4.9 | 1.5 | Iris-versicolor |
| 53 | 5.5 | 2.3 | 4.0 | 1.3 | Iris-versicolor |
| 54 | 6.5 | 2.8 | 4.6 | 1.5 | Iris-versicolor |
| 55 | 5.7 | 2.8 | 4.5 | 1.3 | Iris-versicolor |
| 100 | 6.3 | 3.3 | 6.0 | 2.5 | Iris-virginica |
| 101 | 5.8 | 2.7 | 5.1 | 1.9 | Iris-virginica |
| 102 | 7.1 | 3.0 | 5.9 | 2.1 | Iris-virginica |
| 103 | 6.3 | 2.9 | 5.6 | 1.8 | Iris-virginica |
| 104 | 6.5 | 3.0 | 5.8 | 2.2 | Iris-virginica |
| 105 | 7.6 | 3.0 | 6.6 | 2.1 | Iris-virginica |

Fonte: Kaggle

# / CLASSIFICAÇÃO OU REGRESSÃO?

| | comprim_sepala | largura_sepala | comprim_petala | largura_petala | classe |
|---|---|---|---|---|---|
| 0 | 5.1 | 3.5 | 1.4 | 0.2 | Iris-setosa |
| 1 | 4.9 | 3.0 | 1.4 | 0.2 | Iris-setosa |
| 2 | 4.7 | 3.2 | 1.3 | 0.2 | Iris-setosa |
| 3 | 4.6 | 3.1 | 1.5 | 0.2 | Iris-setosa |
| 4 | 5.0 | 3.6 | 1.4 | 0.2 | Iris-setosa |
| 5 | 5.4 | 3.9 | 1.7 | 0.4 | Iris-setosa |
| 50 | 7.0 | 3.2 | 4.7 | 1.4 | Iris-versicolor |
| 51 | 6.4 | 3.2 | 4.5 | 1.5 | Iris-versicolor |
| 52 | 6.9 | 3.1 | 4.9 | 1.5 | Iris-versicolor |
| 53 | 5.5 | 2.3 | 4.0 | 1.3 | Iris-versicolor |
| 54 | 6.5 | 2.8 | 4.6 | 1.5 | Iris-versicolor |
| 55 | 5.7 | 2.8 | 4.5 | 1.3 | Iris-versicolor |
| 100 | 6.3 | 3.3 | 6.0 | 2.5 | Iris-virginica |
| 101 | 5.8 | 2.7 | 5.1 | 1.9 | Iris-virginica |
| 102 | 7.1 | 3.0 | 5.9 | 2.1 | Iris-virginica |
| 103 | 6.3 | 2.9 | 5.6 | 1.8 | Iris-virginica |
| 104 | 6.5 | 3.0 | 5.8 | 2.2 | Iris-virginica |
| 105 | 7.6 | 3.0 | 6.6 | 2.1 | Iris-virginica |

Fonte: Kaggle

# / CAMPOS DA PLANILHA

| | age | sex | chest pain type | resting bp s | cholesterol | fasting blood sugar | resting ecg | max heart rate | exercise angina | oldpeak | ST slope | target |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 40 | 1 | 2 | 140 | 289 | 0 | 0 | 172 | 0 | 0.0 | 1 | 0 |
| 1 | 49 | 0 | 3 | 160 | 180 | 0 | 0 | 156 | 0 | 1.0 | 2 | 1 |
| 2 | 37 | 1 | 2 | 130 | 283 | 0 | 1 | 98 | 0 | 0.0 | 1 | 0 |
| 3 | 48 | 0 | 4 | 138 | 214 | 0 | 0 | 108 | 1 | 1.5 | 2 | 1 |
| 4 | 54 | 1 | 3 | 150 | 195 | 0 | 0 | 122 | 0 | 0.0 | 1 | 0 |
| 5 | 39 | 1 | 3 | 120 | 339 | 0 | 0 | 170 | 0 | 0.0 | 1 | 0 |
| 6 | 45 | 0 | 2 | 130 | 237 | 0 | 0 | 170 | 0 | 0.0 | 1 | 0 |
| 7 | 54 | 1 | 2 | 110 | 208 | 0 | 0 | 142 | 0 | 0.0 | 1 | 0 |
| 8 | 37 | 1 | 4 | 140 | 207 | 0 | 0 | 130 | 1 | 1.5 | 2 | 1 |
| 9 | 48 | 0 | 2 | 120 | 284 | 0 | 0 | 120 | 0 | 0.0 | 1 | 0 |

COLUNAS

LINHAS

| | age | sex | chest pain type | resting bp s | cholesterol | fasting blood sugar | resting ecg | max heart rate | exercise angina | oldpeak | ST slope | target |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 40 | 1 | 2 | 140 | 289 | 0 | 0 | 172 | 0 | 0.0 | 1 | 0 |
| 1 | 49 | 0 | 3 | 160 | 180 | 0 | 0 | 156 | 0 | 1.0 | 2 | 1 |
| 2 | 37 | 1 | 2 | 130 | 283 | 0 | 1 | 98 | 0 | 0.0 | 1 | 0 |
| 3 | 48 | 0 | 4 | 138 | 214 | 0 | 0 | 108 | 1 | 1.5 | 2 | 1 |
| 4 | 54 | 1 | 3 | 150 | 195 | 0 | 0 | 122 | 0 | 0.0 | 1 | 0 |
| 5 | 39 | 1 | 3 | 120 | 339 | 0 | 0 | 170 | 0 | 0.0 | 1 | 0 |
| 6 | 45 | 0 | 2 | 130 | 237 | 0 | 0 | 170 | 0 | 0.0 | 1 | 0 |
| 7 | 54 | 1 | 2 | 110 | 208 | 0 | 0 | 142 | 0 | 0.0 | 1 | 0 |
| 8 | 37 | 1 | 4 | 140 | 207 | 0 | 0 | 130 | 1 | 1.5 | 2 | 1 |
| 9 | 48 | 0 | 2 | 120 | 284 | 0 | 0 | 120 | 0 | 0.0 | 1 | 0 |

# ATRIBUTOS/FEATURES

| | age | sex | chest pain type | resting bp s | cholesterol | fasting blood sugar | resting ecg | max heart rate | exercise angina | oldpeak | ST slope | target |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 40 | 1 | 2 | 140 | 289 | 0 | 0 | 172 | 0 | 0.0 | 1 | 0 |
| 1 | 49 | 0 | 3 | 160 | 180 | 0 | 0 | 156 | 0 | 1.0 | 2 | 1 |
| 2 | 37 | 1 | 2 | 130 | 283 | 0 | 1 | 98 | 0 | 0.0 | 1 | 0 |
| 3 | 48 | 0 | 4 | 138 | 214 | 0 | 0 | 108 | 1 | 1.5 | 2 | 1 |
| 4 | 54 | 1 | 3 | 150 | 195 | 0 | 0 | 122 | 0 | 0.0 | 1 | 0 |
| 5 | 39 | 1 | 3 | 120 | 339 | 0 | 0 | 170 | 0 | 0.0 | 1 | 0 |
| 6 | 45 | 0 | 2 | 130 | 237 | 0 | 0 | 170 | 0 | 0.0 | 1 | 0 |
| 7 | 54 | 1 | 2 | 110 | 208 | 0 | 0 | 142 | 0 | 0.0 | 1 | 0 |
| 8 | 37 | 1 | 4 | 140 | 207 | 0 | 0 | 130 | 1 | 1.5 | 2 | 1 |
| 9 | 48 | 0 | 2 | 120 | 284 | 0 | 0 | 120 | 0 | 0.0 | 1 | 0 |

# NOME DOS ATRIBUTOS

| | age | sex | chest pain type | resting bp s | cholesterol | fasting blood sugar | resting ecg | max heart rate | exercise angina | oldpeak | ST slope | target |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 40 | 1 | 2 | 140 | 289 | 0 | 0 | 172 | 0 | 0.0 | 1 | 0 |
| 1 | 49 | 0 | 3 | 160 | 180 | 0 | 0 | 156 | 0 | 1.0 | 2 | 1 |
| 2 | 37 | 1 | 2 | 130 | 283 | 0 | 1 | 98 | 0 | 0.0 | 1 | 0 |
| 3 | 48 | 0 | 4 | 138 | 214 | 0 | 0 | 108 | 1 | 1.5 | 2 | 1 |
| 4 | 54 | 1 | 3 | 150 | 195 | 0 | 0 | 122 | 0 | 0.0 | 1 | 0 |
| 5 | 39 | 1 | 3 | 120 | 339 | 0 | 0 | 170 | 0 | 0.0 | 1 | 0 |
| 6 | 45 | 0 | 2 | 130 | 237 | 0 | 0 | 170 | 0 | 0.0 | 1 | 0 |
| 7 | 54 | 1 | 2 | 110 | 208 | 0 | 0 | 142 | 0 | 0.0 | 1 | 0 |
| 8 | 37 | 1 | 4 | 140 | 207 | 0 | 0 | 130 | 1 | 1.5 | 2 | 1 |
| 9 | 48 | 0 | 2 | 120 | 284 | 0 | 0 | 120 | 0 | 0.0 | 1 | 0 |

# CLASSE/RÓTULO

| | age | sex | chest pain type | resting bp s | cholesterol | fasting blood sugar | resting ecg | max heart rate | exercise angina | oldpeak | ST slope | target |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 40 | 1 | 2 | 140 | 289 | 0 | 0 | 172 | 0 | 0.0 | 1 | 0 |
| 1 | 49 | 0 | 3 | 160 | 180 | 0 | 0 | 156 | 0 | 1.0 | 2 | 1 |
| 2 | 37 | 1 | 2 | 130 | 283 | 0 | 1 | 98 | 0 | 0.0 | 1 | 0 |
| 3 | 48 | 0 | 4 | 138 | 214 | 0 | 0 | 108 | 1 | 1.5 | 2 | 1 |
| 4 | 54 | 1 | 3 | 150 | 195 | 0 | 0 | 122 | 0 | 0.0 | 1 | 0 |
| 5 | 39 | 1 | 3 | 120 | 339 | 0 | 0 | 170 | 0 | 0.0 | 1 | 0 |
| 6 | 45 | 0 | 2 | 130 | 237 | 0 | 0 | 170 | 0 | 0.0 | 1 | 0 |
| 7 | 54 | 1 | 2 | 110 | 208 | 0 | 0 | 142 | 0 | 0.0 | 1 | 0 |
| 8 | 37 | 1 | 4 | 140 | 207 | 0 | 0 | 130 | 1 | 1.5 | 2 | 1 |
| 9 | 48 | 0 | 2 | 120 | 284 | 0 | 0 | 120 | 0 | 0.0 | 1 | 0 |

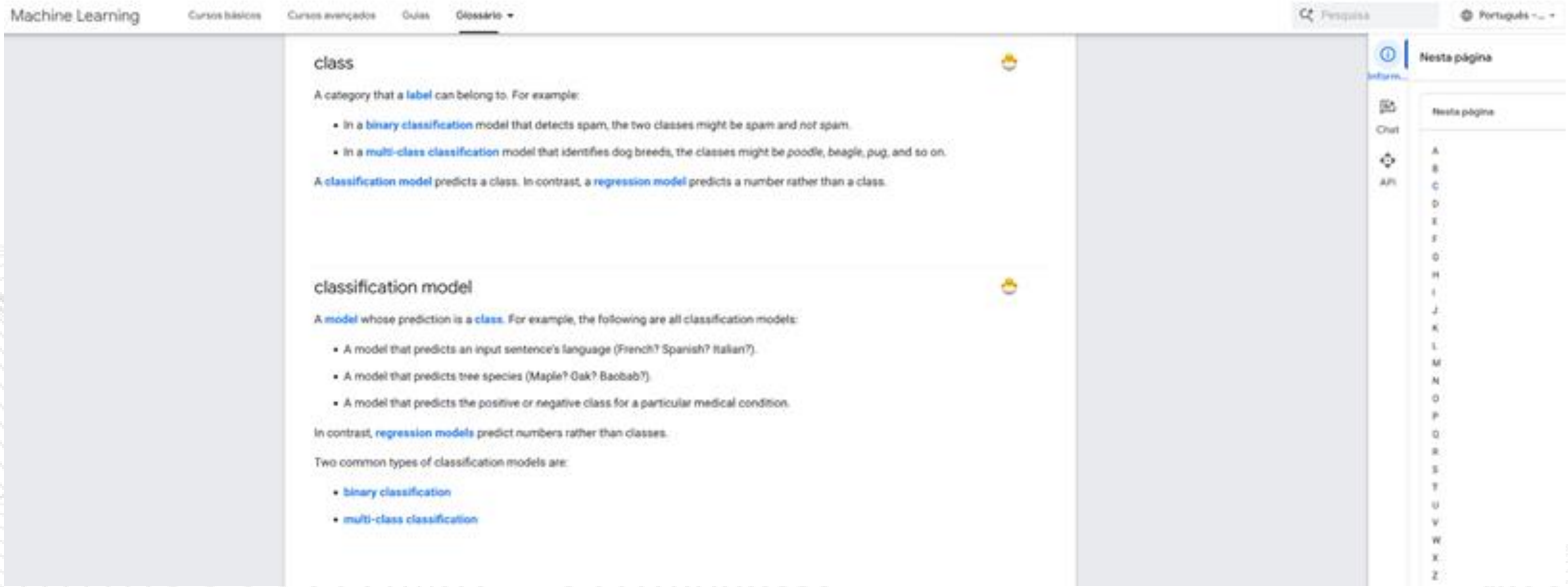
EXEMPLOS OU OBSERVAÇÕES :

| | age | sex | chest pain type | resting bp s | cholesterol | fasting blood sugar | resting ecg | max heart rate | exercise angina | oldpeak | ST slope | target |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 40 | 1 | 2 | 140 | 289 | 0 | 0 | 172 | 0 | 0.0 | 1 | 0 |
| 1 | 49 | 0 | 3 | 160 | 180 | 0 | 0 | 156 | 0 | 1.0 | 2 | 1 |
| 2 | 37 | 1 | 2 | 130 | 283 | 0 | 1 | 98 | 0 | 0.0 | 1 | 0 |
| 3 | 48 | 0 | 4 | 138 | 214 | 0 | 0 | 108 | 1 | 1.5 | 2 | 1 |
| 4 | 54 | 1 | 3 | 150 | 195 | 0 | 0 | 122 | 0 | 0.0 | 1 | 0 |
| 5 | 39 | 1 | 3 | 120 | 339 | 0 | 0 | 170 | 0 | 0.0 | 1 | 0 |
| 6 | 45 | 0 | 2 | 130 | 237 | 0 | 0 | 170 | 0 | 0.0 | 1 | 0 |
| 7 | 54 | 1 | 2 | 110 | 208 | 0 | 0 | 142 | 0 | 0.0 | 1 | 0 |
| 8 | 37 | 1 | 4 | 140 | 207 | 0 | 0 | 130 | 1 | 1.5 | 2 | 1 |
| 9 | 48 | 0 | 2 | 120 | 284 | 0 | 0 | 120 | 0 | 0.0 | 1 | 0 |

# ÍNDICE (INDEX)

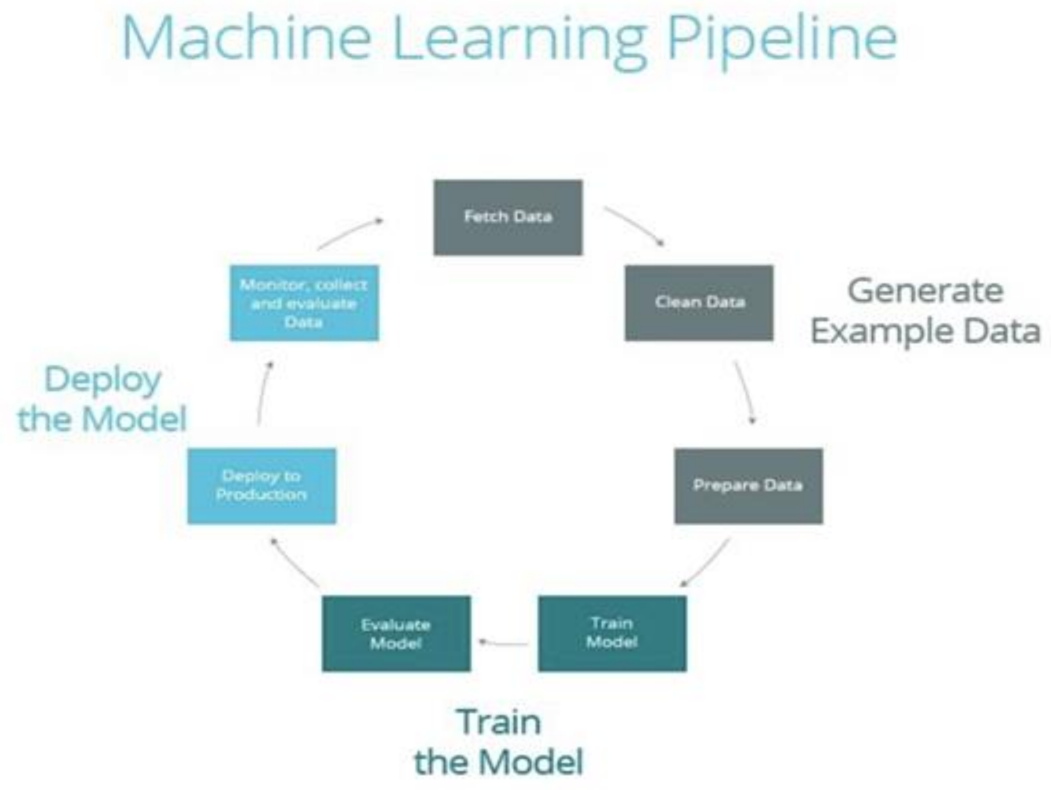
| | age | sex | chest pain type | resting bp s | cholesterol | fasting blood sugar | resting ecg | max heart rate | exercise angina | oldpeak | ST slope | target |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 40 | 1 | 2 | 140 | 289 | 0 | 0 | 172 | 0 | 0.0 | 1 | 0 |
| 1 | 49 | 0 | 3 | 160 | 180 | 0 | 0 | 156 | 0 | 1.0 | 2 | 1 |
| 2 | 37 | 1 | 2 | 130 | 283 | 0 | 1 | 98 | 0 | 0.0 | 1 | 0 |
| 3 | 48 | 0 | 4 | 138 | 214 | 0 | 0 | 108 | 1 | 1.5 | 2 | 1 |
| 4 | 54 | 1 | 3 | 150 | 195 | 0 | 0 | 122 | 0 | 0.0 | 1 | 0 |
| 5 | 39 | 1 | 3 | 120 | 339 | 0 | 0 | 170 | 0 | 0.0 | 1 | 0 |
| 6 | 45 | 0 | 2 | 130 | 237 | 0 | 0 | 170 | 0 | 0.0 | 1 | 0 |
| 7 | 54 | 1 | 2 | 110 | 208 | 0 | 0 | 142 | 0 | 0.0 | 1 | 0 |
| 8 | 37 | 1 | 4 | 140 | 207 | 0 | 0 | 130 | 1 | 1.5 | 2 | 1 |
| 9 | 48 | 0 | 2 | 120 | 284 | 0 | 0 | 120 | 0 | 0.0 | 1 | 0 |

# / GLOSSÁRIO




FONTE: GOOGLE

# / ML PIPELINE

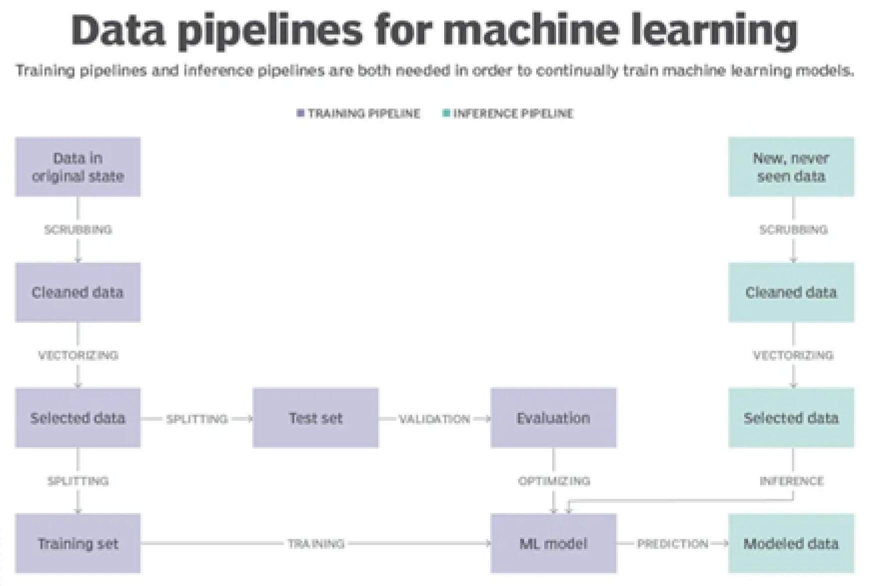



FONTE: DATATRON

# ETAPAS TREINAMENTO E INFERÊNCIA

/01

O conjunto de dados é **dividido** em treino e teste.

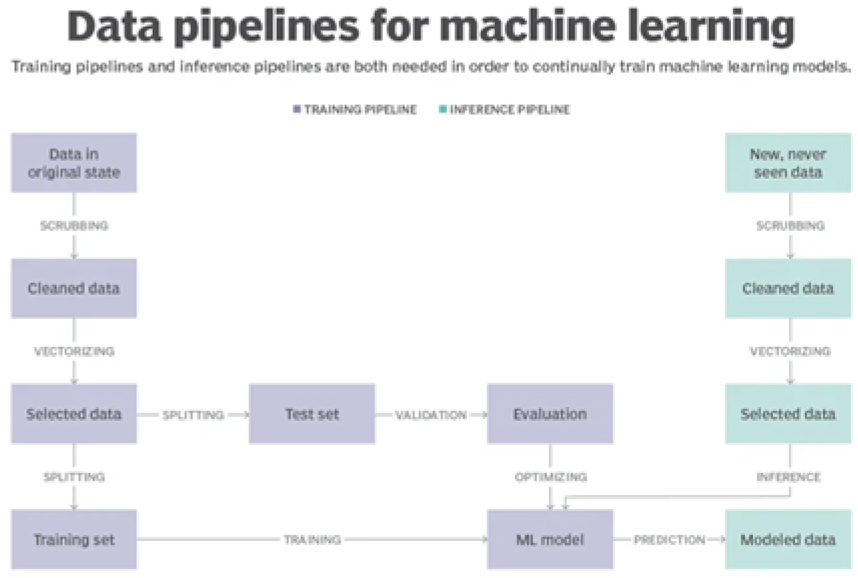



FONTE: AWS SAGE MAKER

# ETAPAS TREINAMENTO E INFERÊNCIA

## /02

Aplicação do algoritmo de aprendizado de máquina no conjunto de treino. Isso fará com que ele **aprenda** a relação entre os **atributos** e a respectiva **classe** para cada exemplo.

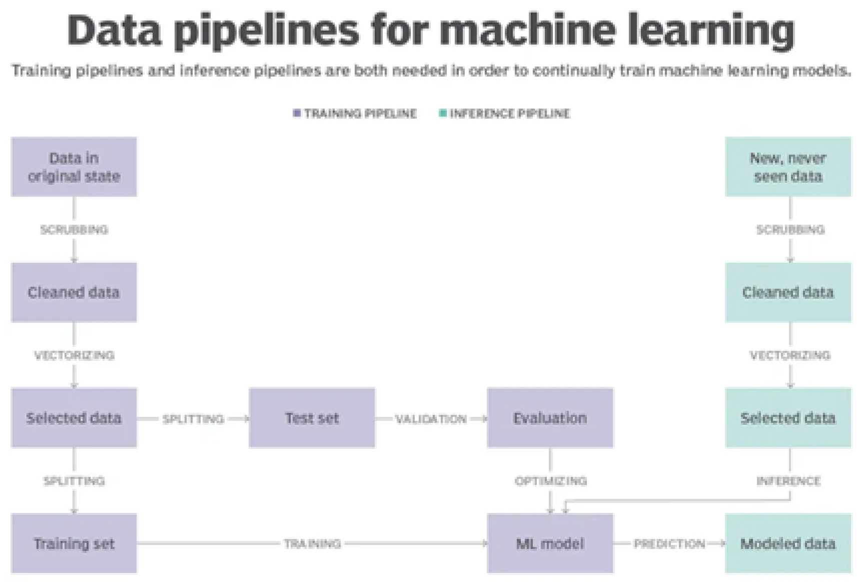



FONTE: AWS SAGE MAKER

# ETAPAS TREINAMENTO E INFERÊNCIA

## /03

Com o algoritmo treinado, mostra-se para ele apenas os **atributos** do conjunto de teste para obter as respostas das **classes** que o algoritmo acredita ser para cada exemplo baseado em seu aprendizado. Neste momento, a coluna com as classes são omitidas para o algoritmo.

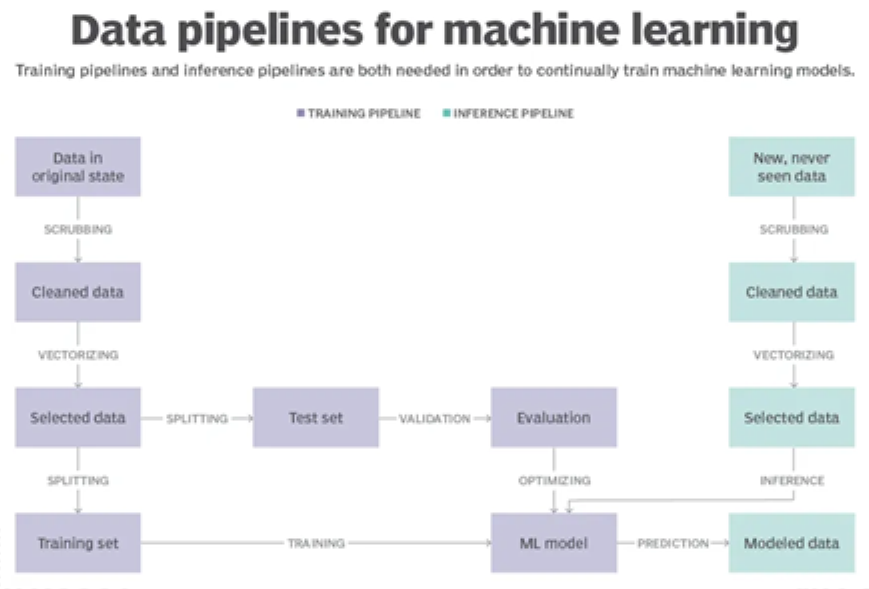



FONTE: AWS SAGE MAKER

# ETAPAS TREINAMENTO E INFERÊNCIA

## /04

**Comparamos** as respostas obtidas das classes pelo algoritmo com as **verdadeiras** e utilizamos alguma métrica para avaliar o desempenho.

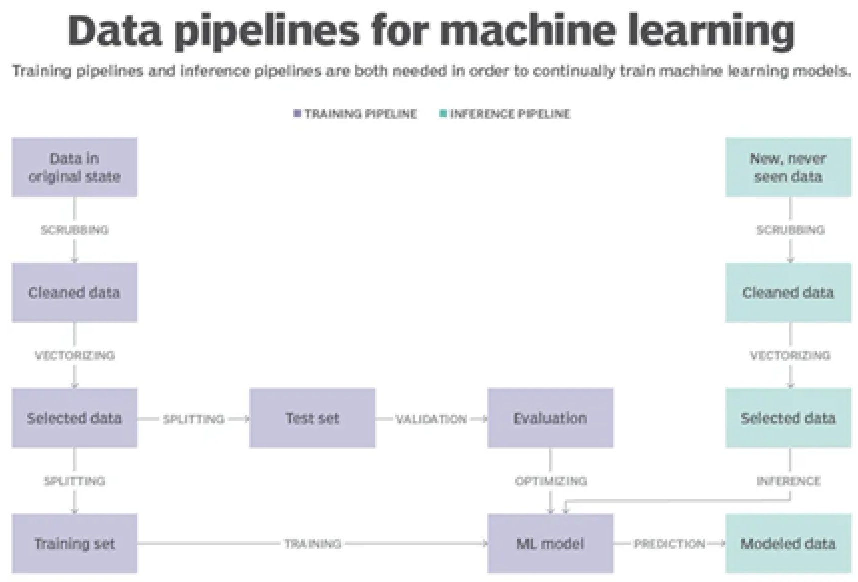



FONTE: AWS SAGE MAKER

# ETAPAS TREINAMENTO E INFERÊNCIA

## /05

Salvamos o resultado da **métrica**, voltamos na etapa inicial e realizamos ajustes para tentar melhorar o desempenho do algoritmo.

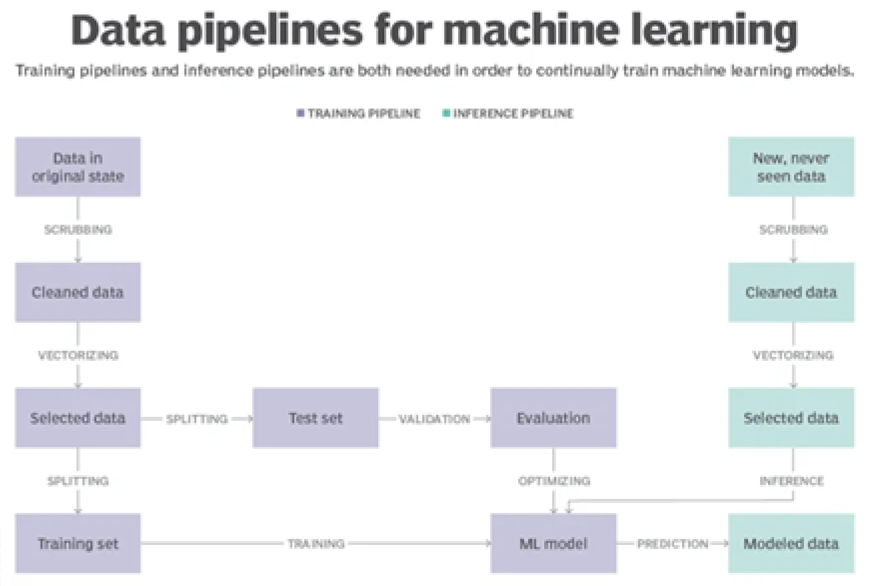



FONTE: AWS SAGE MAKER

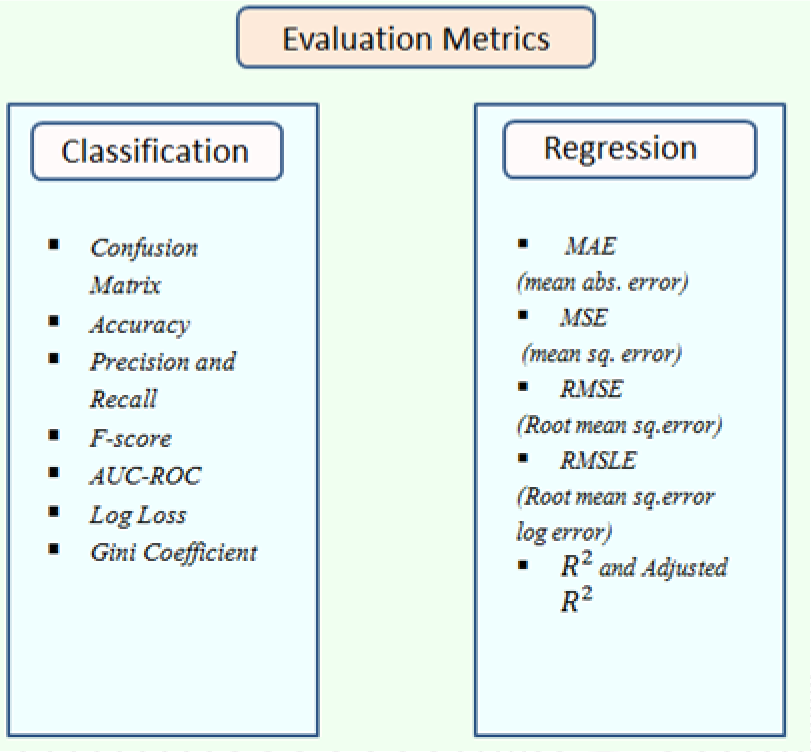
# ETAPAS TREINAMENTO E INFERÊNCIA

## /06

Com os ajustes feitos, temos um **modelo** de aprendizado de máquina.




FONTE: AWS SAGE MAKER




FONTE: ML EVALUATION METRICS

# MÉTRICAS DE AVALIAÇÃO

## Classificação

- Acurácia
- Precisão
- Revocação
- F1-Score
- Matriz de confusão

## Regressão

- Erro quadrático médio (MSE)
- Raiz do erro quadrático médio (RMSE)
- Erro percentual absoluto médio (MAPE)

POS TECH

FIAP + alura