

• Os <u>frameworks de deep learning</u> são bibliotecas de software que fornecem ferramentas e recursos para construir e treinar modelos de deep learning, tornando o processo mais eficiente e acessível.

• TensorFlow:

- Características:

- Flexível: Pode ser usado para tarefas complexas, desde modelos de visão computacional até processamento de linguagem natural.
- Escalabilidade: Suporta treinamento em GPUs e TPUs para acelerar o processo.
- Comunidade ativa: Grande comunidade de usuários, com amplo suporte e recursos online.
- Produção pronta: Ferramentas para implementar modelos em produção, como TensorFlow Serving.

– Exemplos de uso:

- Reconhecimento de imagens (Google Photos, carros autônomos)
- Classificação de texto (filtragem de spam, tradução automática)
- Previsão (previsão de séries temporais, análise de mercado)





PyTorch:

- Características:

- Facilidade de uso: Interface intuitiva e amigável para desenvolvedores.
- **Dinamicidade:** Permite a definição de gráficos computacionais dinâmicos, tornando-o ideal para pesquisa.
- Comunidade vibrante: Grande comunidade de desenvolvedores, com fóruns ativos e recursos de aprendizado.
- **Popular na pesquisa:** Largamente utilizado na comunidade acadêmica devido à sua flexibilidade.

– Exemplos de uso:

- Processamento de linguagem natural (chatbots, modelos de linguagem)
- Visão computacional (segmentação de imagens, detecção de objetos)
- Redes neurais generativas (GANs)



· <u>Keras:</u>

- Características:
 - Simples e intuitivo: API de alto nível que simplifica o desenvolvimento de modelos.
 - Facilidade de uso: Fácil de aprender e implementar, ideal para iniciantes.
 - Modular: Permite a criação de modelos complexos combinando camadas e funções.
 - Suporte para vários backends: Pode ser usado com TensorFlow, Theano e CNTK.
- Exemplos de uso:
 - Classificação de imagens (classificação de dígitos, reconhecimento facial)
 - Previsão de séries temporais (previsão de preços de ações)
 - Sistemas de recomendação (recomendação de produtos em lojas online)

· <u>Theano:</u>

- Características:
 - Eficiência computacional: Permite a otimização automática de código para GPUs.
 - Definido por grafo: Permite a definição de modelos através de grafos computacionais.
 - Flexibilidade: Permite a criação de modelos personalizados e inovadores.
 - Foco em pesquisa: Utilizado em pesquisas acadêmicas, embora não seja tão popular para desenvolvimento.
- Exemplos de uso:
 - Modelos de visão computacional (reconhecimento de objetos, detecção de faces)
 - Modelos de processamento de linguagem natural (tradução automática, análise de sentimentos)
 - Redes neurais recorrentes (RNNs)

MXNet:

– Características:

- Desempenho e escalabilidade: Permite o treinamento em GPUs e CPUs, com escalabilidade para clusters.
- Eficiência de memória: Utiliza técnicas de otimização para minimizar o uso de memória.
- Suporte a múltiplas linguagens: Permite o desenvolvimento em Python, C++, R, Julia e JavaScript.
- Comunidade ativa: Grande comunidade de usuários, com recursos e documentação abrangentes.

– Exemplos de uso:

- Visão computacional (reconhecimento de objetos, segmentação de imagens)
- · Processamento de linguagem natural (tradução automática, análise de sentimentos)
- · Sistemas de recomendação (recomendação de música e filmes)

• <u>CNTK:</u>

- Características:
 - Alto desempenho: Projetado para treinamentos de larga escala e execução eficiente.
 - Escalabilidade: Suporta treinamento em múltiplas GPUs e clusters.
 - Orientado a grafos: Permite a definição de modelos complexos através de grafos computacionais.
 - Linguagens de programação: Suporta C++, Python e C#.
- Exemplos de uso:
 - Reconhecimento de fala (assistentes virtuais, tradução de voz)
 - Processamento de linguagem natural (classificação de texto, análise de sentimentos)
 - Modelos de previsão (previsão de demanda, análise de riscos)





A escolha do framework depende da necessidade específica do projeto.

Considerar fatores como facilidade de uso, performance, escalabilidade e comunidade é crucial.

Cada framework oferece recursos únicos e vantagens distintas, permitindo a escolha ideal para o desenvolvimento de modelos de deep learning eficientes e inovadores.

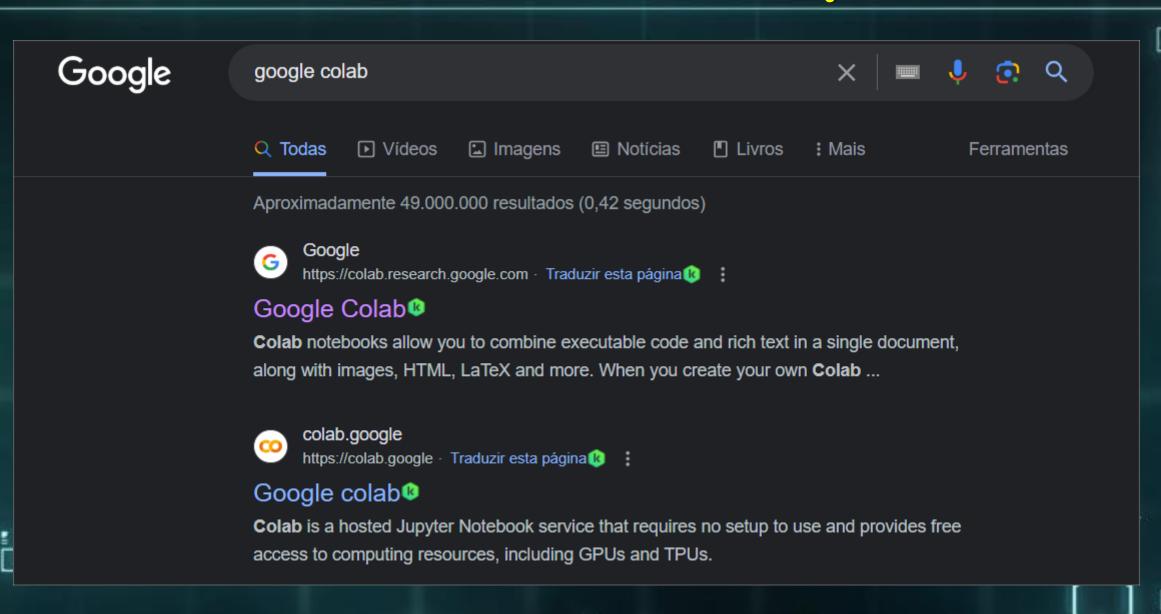
Ambiente de programação

- Editor de Python
- Colab do Google, pesquise no próprio Google search

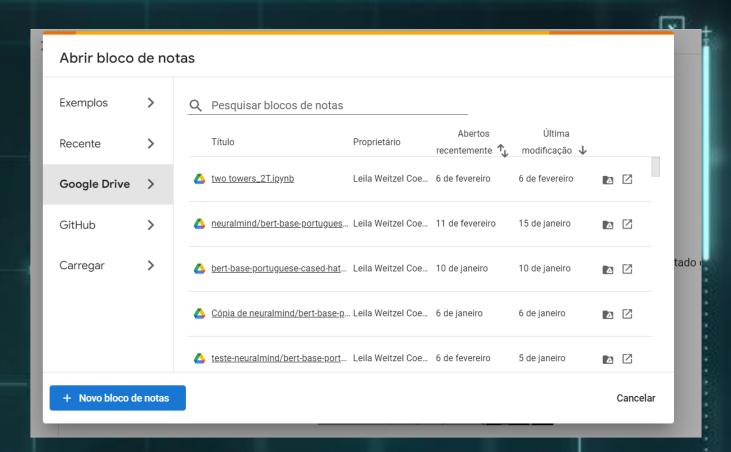


Ambiente de programação

[__]



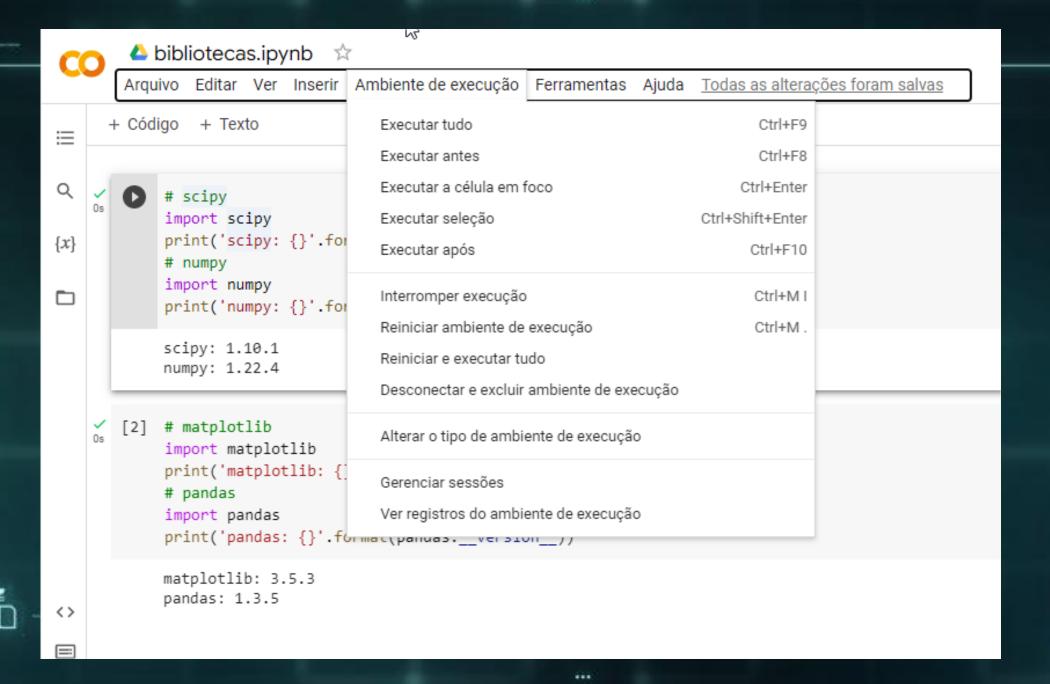
- Abrir seu Google Drive e abrir os exemplos que foram dados.
- Todos os Arquivos são
 disponibilizados no Classroom,
 no drive da disciplina
- Vocês devem fazer uma cópia para o drive de vocês dos exercícios e exemplos dados



Dúvidas ler tutorial

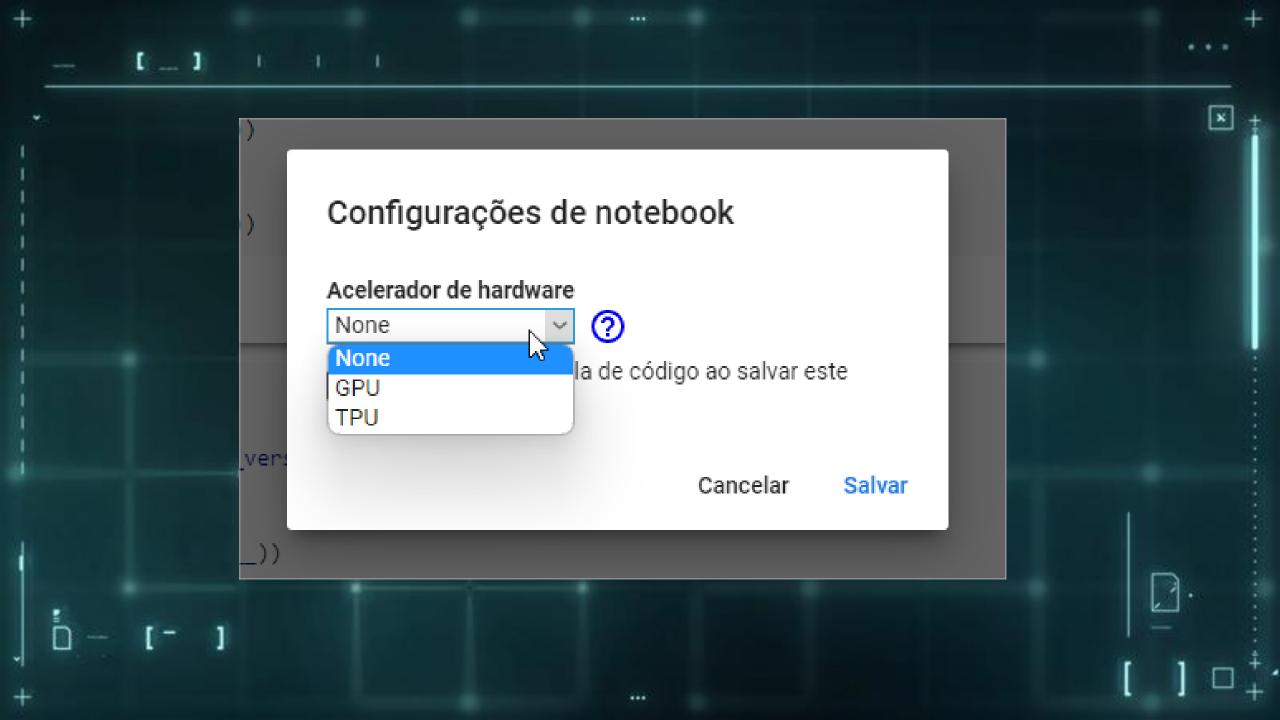
https://www.alura.com.br/artigos/google-colab-o-que-e-e-como-usar

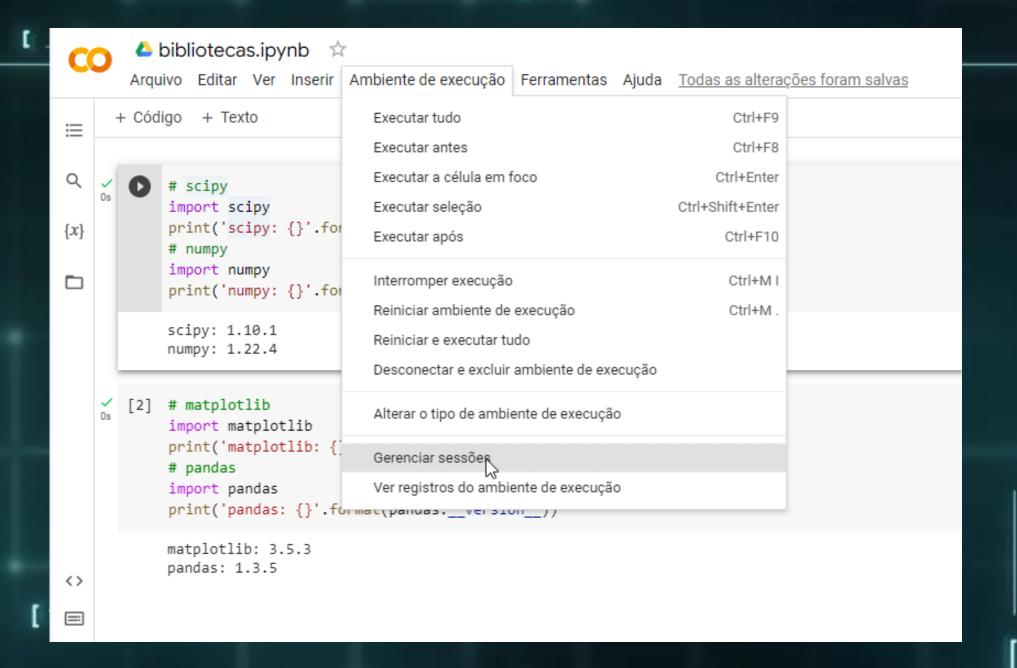




	· inv	nh 🕾	,							
[]	s.ipynb 😭									
	Ver	Inserir	Ambiente de execução	Todas as altera	ç <u>ões fora</u> ı					
)		Executar tudo	Executar tudo						
			Executar antes			Ctrl+F8				
			Executar a célula em f	Ctrl+Enter						
	py py: {]		Executar seleção	Ctrl+Shift+Enter						
		{}'.for	Executar após	Ctrl+F10						
	py py:	{}'.for	Interromper execução	Interromper execução						
		Reiniciar ambiente de	Ctrl+M .							
	0.1 2.4		Reiniciar e executar tu							
	2.4	_	Desconectar e excluir							
	ib plotlib		Alterar o tipo de ambie	ente de execuçã	0					
	plot	lib: {]	Gerenciar sessões							
	das		Ver registros do ambie	Ver registros do ambiente de execução						
[-	das:	{}'.fo	umac(hannasversic							
	: 3. 3.5	5.3								

×





Sessões ativas

Título Última execução RAM usada

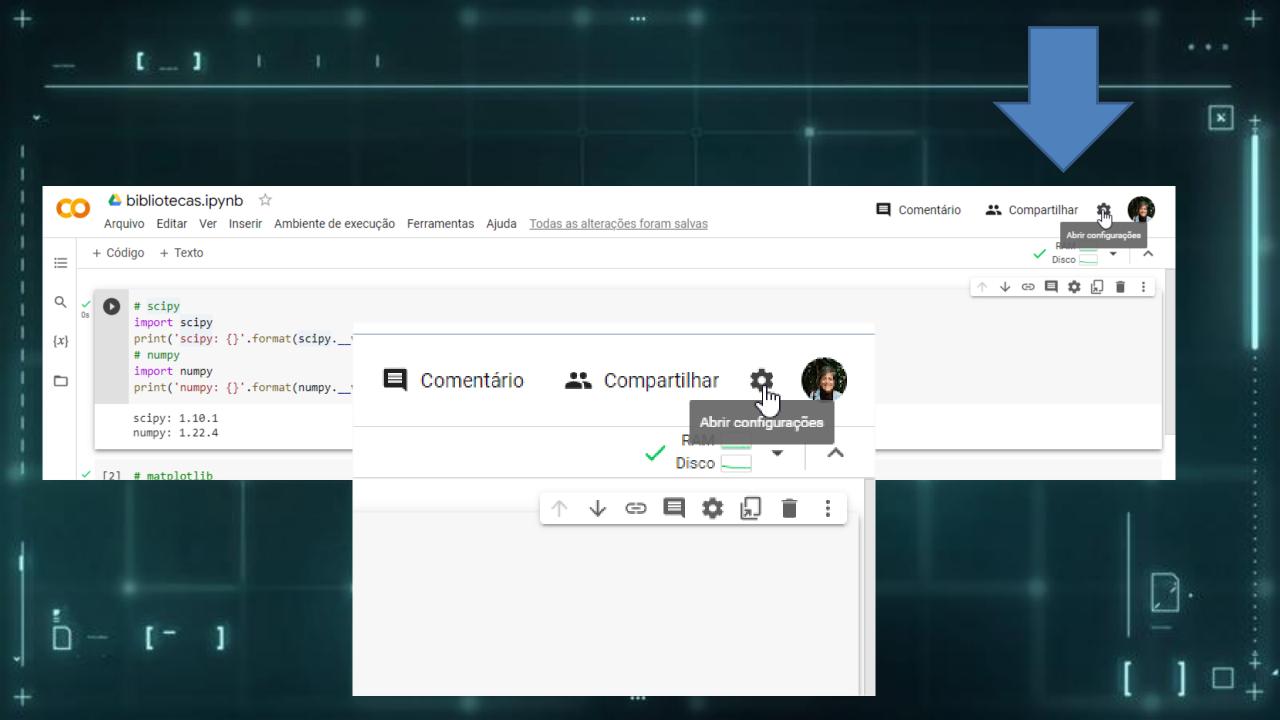


bibliotecas.ipynb Sessão atual

há 0 minuto

0.89 GB

ENCERRAR



Configurações

 Site
 Tema adaptive
 ▼

 Editor
 ✓ Exibir notificações na área de trabalho quando as execuções forem concluídas ☐ Os notebooks novos usam saídas particulares (omitem saídas ao salvar)

 Colab Pro
 Layout padrão da página horizontal
 ▼

 GitHub
 ✓

 Diversos
 Custom snippet notebook URL

 ☐ Usar um notebook temporário como página de destino padrão.

Cancelar

Salvar

Configurações

Site

Conta do GitHub

Editor

Sua conta do GitHub está conectada com sua conta do Google Drive.

Colab Pro

Revogar acesso em github.com

Esquecer conta do GitHub

GitHub

Acessar repositórios e organizações particulares Mais informações

Diversos

L,

• Tupla: leitura apenas

[_]



• Lista: As listas usam a notação de colchete e podem ser indexadas usando a notação de matriz

[]

```
mylist = [1, 2, 3]
print("Zeroth Value:", mylist[0])
mylist.append(4)
print("List Length:", len(mylist))
for value in mylist:
 print (value)
Zeroth Value: 1
List Length: 4
```

[]

• Dicionário: Os dicionários são mapeamentos como pares de valores-chave.

```
mydict = {'a': 1, 'b': 2, 'c': 3}
print("A value: ", mydict['a'])
mydict['a'] = 11
print("A value:", mydict['a'])
print("Keys:", mydict.keys())
print("Values:",mydict.values())
for key in mydict.keys():
 print (mydict[key])
A value: 1
A value: 11
Keys: dict_keys(['a', 'b', 'c'])
Values: dict_values([11, 2, 3])
```

• função

```
# Sum function

def mysum(x, y):

    return x + y

# Test sum function

result = mysum(1, 3)

print(result)

4
```

Numpy

```
import numpy
mylist = [[1, 2, 3], [3, 4, 5]]
myarray = numpy.array(mylist)
print(myarray)
print(myarray.shape)
print("First row: ",myarray[0])
print("Last row: %s", myarray[-1])
print("Specific row and col: ", myarray[0, 2])
print("Whole col: %s", myarray[:, 2])
[[1 2 3]
 [3 4 5]]
(2, 3)
First row: [1 2 3]
Last row: %s [3 4 5]
Specific row and col: 3
Whole col: %s [3 5]
```

Matplotlib

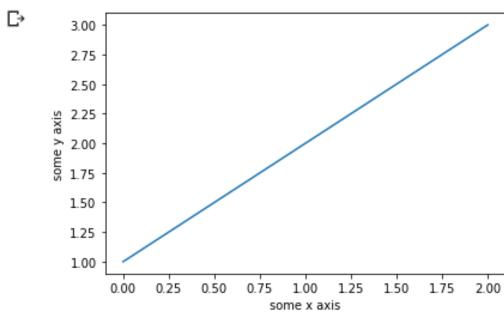
[_]

•

import numpy as np

```
# basic line plot
import matplotlib.pyplot as plt
import numpy
myarray = numpy.array([1, 2, 3])
plt.plot(myarray)
plt.xlabel('some x axis')
plt.ylabel('some y axis')
plt.show()
```

As é um alias = apelido



[] # basic scatter plot Matplotlib import matplotlib.pyplot as plt import numpy x = numpy.array([1, 2, 3])y = numpy.array([2, 4, 6])plt.scatter(x,y) plt.xlabel('some x axis') plt.ylabel('some y axis') plt.show() Ľ→ 6.0 5.5 5.0 3.5 axis 3.0 2.5 2.0 -1.25 2.50 2.75 1.50 1.75 2.00 2.25 3.00 some x axis

Pandas

[_]

- Series
- unidimensional

```
# series
import numpy
import pandas
myarray = numpy.array([1, 2, 3])
rownames = ['a', 'b', 'c']
myseries = pandas.Series(myarray, index=rownames)
print(myseries)
dtype: int64
```



- Pandas
- Dataframe multidimensional

```
# dataframe
import numpy
import pandas
myarray = numpy.array([[1, 2, 3], [4, 5, 6]])
rownames = ['a', 'b']
colnames = ['one', 'two', 'three']
mydataframe = pandas.DataFrame(myarray, index=rownames, columns=colnames)
print(mydataframe)
```

Ler arquivos

- Os arquivos estão sempre no formato **csv** (arquivo texto onde as colunas são separados por virgula)
- Atenção se abrir no Excel este arquivo , automaticamente é trocado a virgula por ponto e virgula... Então não abrir....

Ler arquivos

• 1. Carregar arquivos CSV com a biblioteca padrão Python.

- 2. Carregar arquivos CSV com o NumPy.
- 3. Carregar arquivos CSV com Pandas.



Ler arquiyos

- Os comentários em um arquivo CSV são indicados por um hash (#) no início de uma linha.
- Se você tiver comentários em seu arquivo, dependendo do método utilizado para carregar seu dados, você pode precisar indicar se deve ou não esperar comentários e o caráter a esperar para significar uma linha de comentário.

Ler arquivos

- · Duas formas de leitura de dados:
 - Por meio de leitura do seu próprio dataset.
 - –Por meio de **load** pois tem algumas bibliotecas que já têm dados pré-coletados

```
>
```

```
(28] import pandas as pd

df = pd.read_csv('/content/arquivo1.csv')
```

Lendo um arquivo de upload

(1795, 7)

	id	my_datetime	event	country	user_id	source	topic
0	0	2018-01-01 00:01:01	read	country_7	2458151261	SEO	North America
1	1	2018-01-01 00:03:20	read	country_7	2458151262	SEO	South America
2	2	2018-01-01 00:04:01	read	country_7	2458151263	AdWords	Africa
3	3	2018-01-01 00:04:02	read	country_7	2458151264	AdWords	Europe
4	4	2018-01-01 00:05:03	read	country_8	2458151265	Reddit	North America

[24] import pandas as pd import numpy as np from sklearn.datasets import fetch_california_housing data = fetch_california_housing(as_frame=True)

Lendo um arquivo por load do sklearn

0	data.data	
---	-----------	--

	MedInc	HouseAge	AveRooms	AveBedrms	Population	Ave0ccup	Latitude	Longitude
0	8.3252	41.0	6.984127	1.023810	322.0	2.555556	37.88	-122.23
1	8.3014	21.0	6.238137	0.971880	2401.0	2.109842	37.86	-122.22
2	7.2574	52.0	8.288136	1.073446	496.0	2.802260	37.85	-122.24



```
data.target

0 4.526
1 3.585
2 3.521
```

4 3.422 ... 20635 0.781 20636 0.771 20637 0.923

3.413

20638 0.847 20639 0.894

Name: MedHouseVal, Length: 20640, dtype: float64

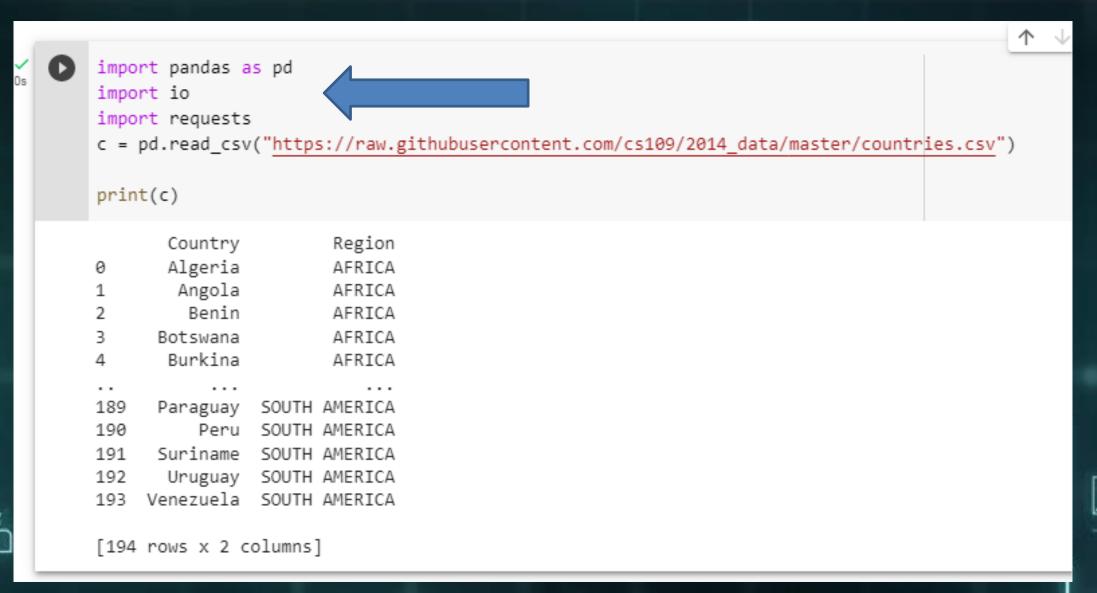
https://scikit-learn.org/stable/modules/classes.html#module-sklearn.datasets

Loaders	
datasets.clear_data_home([data_home])	Delete all the content of the data home cache.
datasets.dump_svmlight_file(X, y, f, *[,])	Dump the dataset in symlight / libsym file format.
datasets.fetch_20newsgroups(*[, data_home,])	Load the filenames and data from the 20 newsgroups dataset (classification)
datasets.fetch_20newsgroups_vectorized(*[,])	Load and vectorize the 20 newsgroups dataset (classification).
datasets.fetch_california_housing(*[,])	Load the California housing dataset (regression).
datasets.fetch_covtype(*[, data_home,])	Load the covertype dataset (classification).
datasets.fetch_kddcup99(*[, subset,])	Load the kddcup99 dataset (classification).
datasets.fetch_lfw_pairs(*[, subset,])	Load the Labeled Faces in the Wild (LFW) pairs dataset (classification).
datasets.fetch_lfw_people(*[, data_home,])	Load the Labeled Faces in the Wild (LFW) people dataset (classification).
datasets.fetch_olivetti_faces(*[,])	Load the Olivetti faces data-set from AT&T (classification).
datasets.fetch_openml([name, version,])	Fetch dataset from openml by name or dataset id.
datasets.fetch_rcv1(*[, data_home, subset,])	Load the RCV1 multilabel dataset (classification).
datasets.fetch_species_distributions(*[,])	Loader for species distribution dataset from Phillips et.
datasets.get_data_home([data_home])	Return the path of the scikit-learn data directory.
datasets.load_breast_cancer(*[, return_X_y,])	Load and return the breast cancer wisconsin dataset (classification).
datasets.load_diabetes(*[, return_X_y,])	Load and return the diabetes dataset (regression).
datasets.load_digits(*[, n_class,])	Load and return the digits dataset (classification).
datasets.load_files(container_path, *[,])	Load text files with categories as subfolder names.
datasets.load_iris(*[, return_X_y, as_frame])	Load and return the iris dataset (classification).
datasets.load_linnerud(*[, return_X_y, as_frame])	Load and return the physical exercise Linnerud dataset.
datasets.load sample image(image name)	Load the numpy array of a single sample image.



Lendo arquivo de url

[]



Lendo outros formatos

- Ler json
- Ler arff
- Ler excel
- Tutorial no link
 - https://practicaldatascience.co.uk/data-science/how-to-import-data-into-pandas-dataframes

Gerar um base de dados

Samples generator

[_]

datasets.make_biclusters(shape, n_clusters, *)	Generate a constant block diagonal structure array for biclustering.
datasets.make_blobs([n_samples, n_features,])	Generate isotropic Gaussian blobs for clustering.
<pre>datasets.make_checkerboard(shape, n_clusters, *)</pre>	Generate an array with block checkerboard structure for biclustering.
datasets.make_circles([n_samples, shuffle,])	Make a large circle containing a smaller circle in 2d.
datasets.make_classification([n_samples,])	Generate a random n-class classification problem.
datasets.make_friedman1([n_samples,])	Generate the "Friedman #1" regression problem.
datasets.make_friedman2([n_samples, noise,])	Generate the "Friedman #2" regression problem.
datasets.make_friedman3([n_samples, noise,])	Generate the "Friedman #3" regression problem.
datasets.make_gaussian_quantiles(*[, mean,])	Generate isotropic Gaussian and label samples by quantile.
datasets.make_hastie_10_2([n_samples,])	Generate data for binary classification used in Hastie et al. 2009, Example 10.2
<pre>datasets.make_low_rank_matrix([n_samples,])</pre>	Generate a mostly low rank matrix with bell-shaped singular values.
datasets.make_moons([n_samples, shuffle,])	Make two interleaving half circles.
<pre>datasets.make_multilabel_classification([])</pre>	Generate a random multilabel classification problem.
<pre>datasets.make_regression([n_samples,])</pre>	Generate a random regression problem.
datasets.make_s_curve([n_samples, noise,])	Generate an S curve dataset.
<pre>datasets.make_sparse_coded_signal(n_samples,)</pre>	Generate a signal as a sparse combination of dictionary elements.
datasets.make_sparse_spd_matrix([dim,])	Generate a sparse symmetric definite positive matrix.
datasets.make_sparse_uncorrelated([])	Generate a random regression problem with sparse uncorrelated design.
datasets.make_spd_matrix(n_dim, *[,])	Generate a random symmetric, positive-definite matrix.
datasets.make_swiss_roll([n_samples, noise,])	Generate a swiss roll dataset.
1	

Gerar um base de dados

```
from sklearn.datasets import make_blobs
X, y = make_blobs(n_samples=10, centers=3, n_features=2,
                  random_state=0)
print(X.shape)
X, y = make_blobs(n_samples=[3, 3, 4], centers=None, n_features=2,
                   random_state=0)
print(X.shape)
(10, 2)
(10, 2)
```

[_]

Gerar um base de dados

```
import pandas as pd
from sklearn.datasets import make_classification
# Gerar um conjunto de dados de classificação
X, y = make_classification(n_samples=1000, n_features=20, n_informative=2,
n_redundant=2, n_classes=2, random_state=42)
# Converter para DataFrame
df = pd.DataFrame(X, columns=[f'feature_{i}' for i in range(X.shape[1])])
df['target'] = y
# Salvar em um arquivo CSV
df.to_csv('classification_data.csv', index=False)
print("Arquivo 'classification_data.csv' criado com sucesso!")
```

Lendo um arquivo csy

• Lendo um arquivo csv que não contem os nomes da coluna.



	preg	plas	pres	skin	test	mass	pedi	age	class	0
0	6	148	72	35	0	33.6	0.627	50	1	
1	1	85	66	29	0	26.6	0.351	31	0	
2	8	183	64	0	0	23.3	0.672	32	1	

Machine Learning

Modelos preditivos apresentam 6 fases

- 1. Definir o problema: a fim de compreender melhor os objetivos do projeto.
- 2. Analisar os dados: Utilizar estatísticas descritivas e visualização para melhor compreender os dados
- 3. Preparar os dados: Use transformadores de dados (por exemplo normalização) para melhor expor a estrutura do problema de predição para algoritmos de modelagem.
- 4. Avaliar Algoritmos: Projetar um bateria de teste para avaliar uma série de algoritmos padrão
- 5. Melhore os resultados: Use métodos de ajuste para obter o máximo de performance dos algoritmos em seus dados
- 6. Apresente os Resultados: Finalize o modelo, faça previsões e apresente os resultados



Machine Learning

• Passos do treinamento:

- 1. Criar o modelo
- 2. Compilar o modelo: decidir quais hiperparametros usar
- 3. Treinar o modelo (função fit()) dividir a base em treinamento (70%) e teste (30%).
- 4. Fazer predições y_pred().
- 5. Avaliar o modelo





See You Next time