

AULA 8

GOOGLE COLABORATORY

1. Objetivos

- Introduzir o ambiente de desenvolvimento Google Colaboratory (Colab).
- Configuração de hardware disponibilizada pelo Colab.
- Notebooks e arquivos no Colab.
- Alguns comandos úteis.

2. Aspectos gerais

- O Google fornece um serviço de nuvem gratuito baseado nos Notebooks Jupyter que suporta GPU.
- No Colab existem limites de uso para tempo e memória \Rightarrow funciona somente durante 12 horas sem interrupção, sendo que após 12 horas ele fecha sozinho e se o trabalho não foi salvo perde-se tudo.
- A configuração de hardware e software disponibilizada pelo Google Colab é a seguinte:
 - CPU com memória RAM de até 48 GB;
 - GPU NVIDIA Tesla K80 de 12 GB de memória;
 - Inúmeras bibliotecas pré-instaladas, tais como, Python, TensorFlow, Scikit-learn, Matplotlib etc.
- O Colab se parece muito com um Notebook Jupyter, assim, o seu uso é fácil para quem já trabalha com essa interface.
- Mais informações sobre o Colab podem ser obtidas em:
<https://colab.research.google.com/notebooks/intro.ipynb#>

3. Iniciar uma sessão e execução de notebooks

- Para usar o Google Colab você tem que ter que possuir um email no gmail.com do Google.

- A forma mais fácil de iniciar o Google Colab é entrar no seu gmail e acessar pela sua conta o Drive, conforma mostrado na Figura 1.

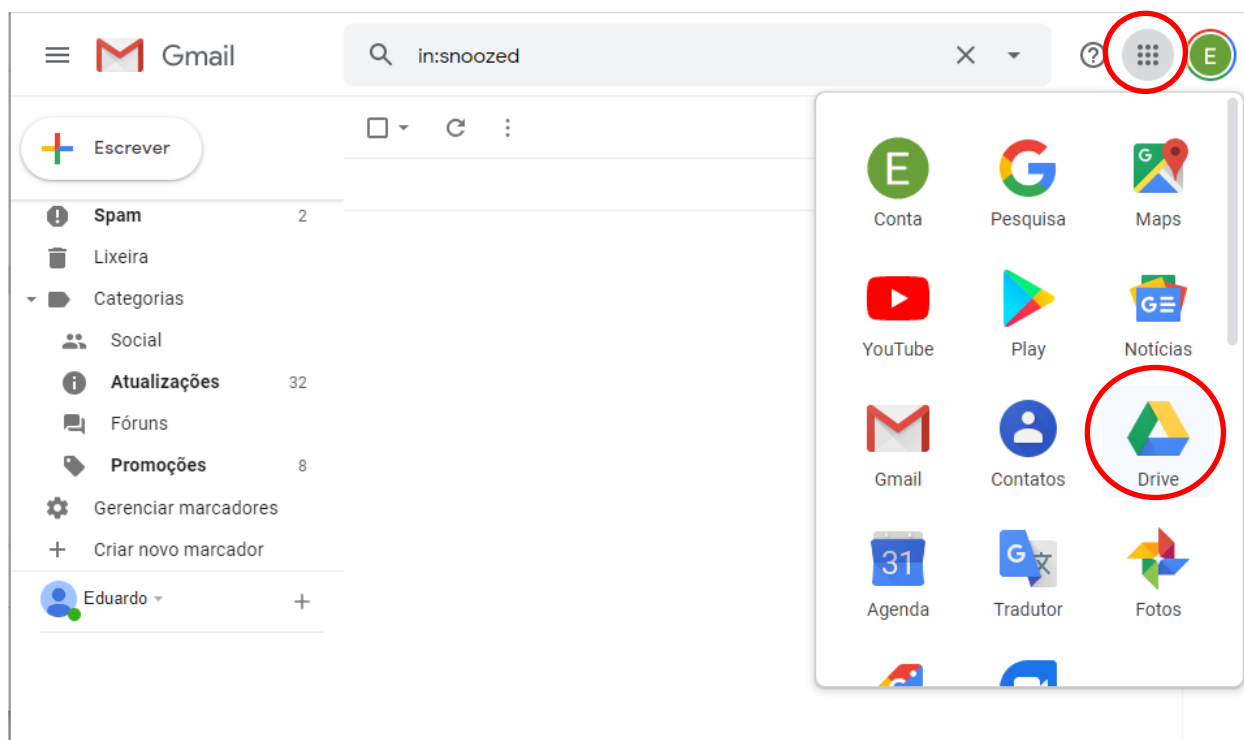


Figura 1. Abrir o Google Drive.

- Para iniciar o Colab, dentro do Google Drive clique com o botão direito do mouse em: Meu Drive > Mais > Google Colaboratory, conforme mostrado na Figura 2.

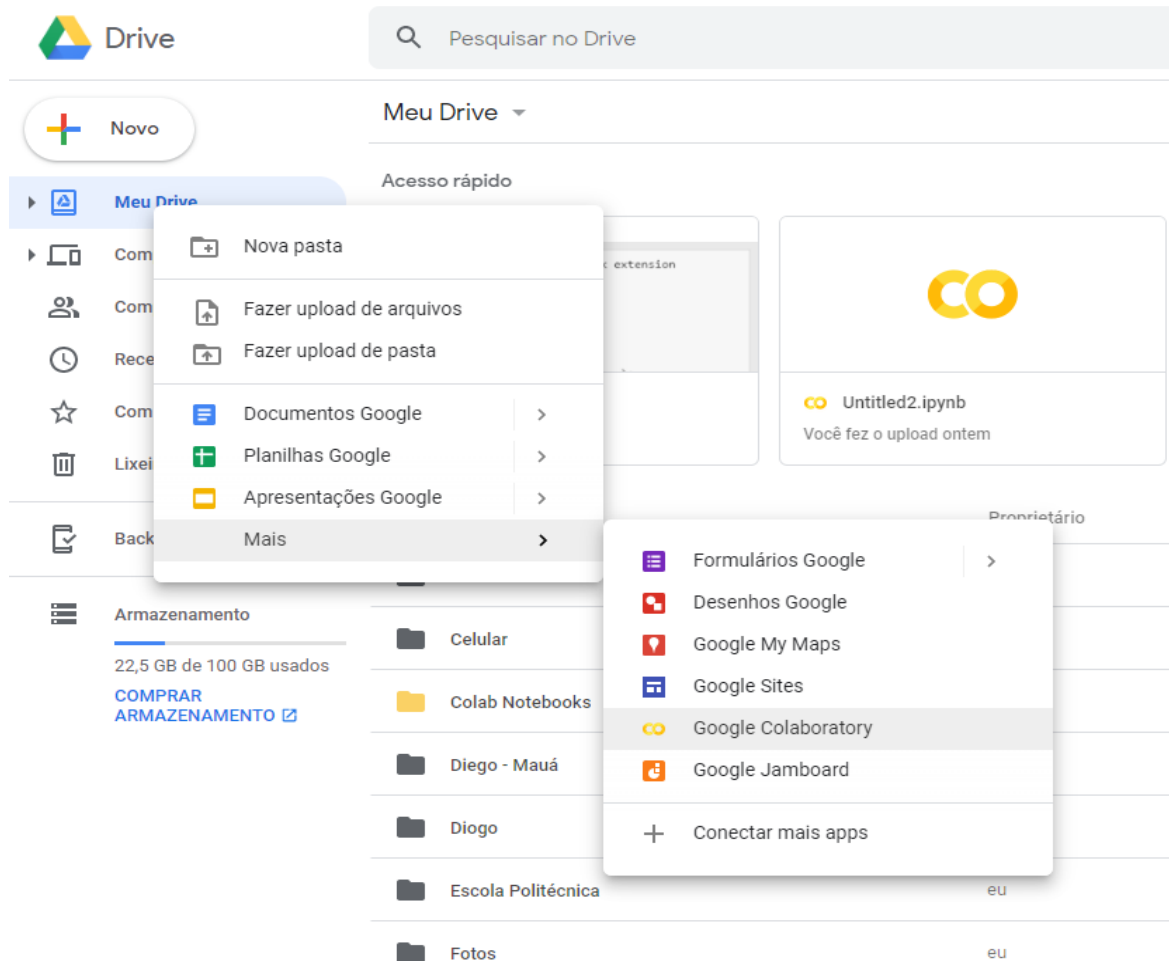


Figura 2. Iniciar o Google Colab.

- Com isso, você inicia um novo documento no Colab, parecido com um Notebook do Jupyter. Veja na Figura 3 abaixo.

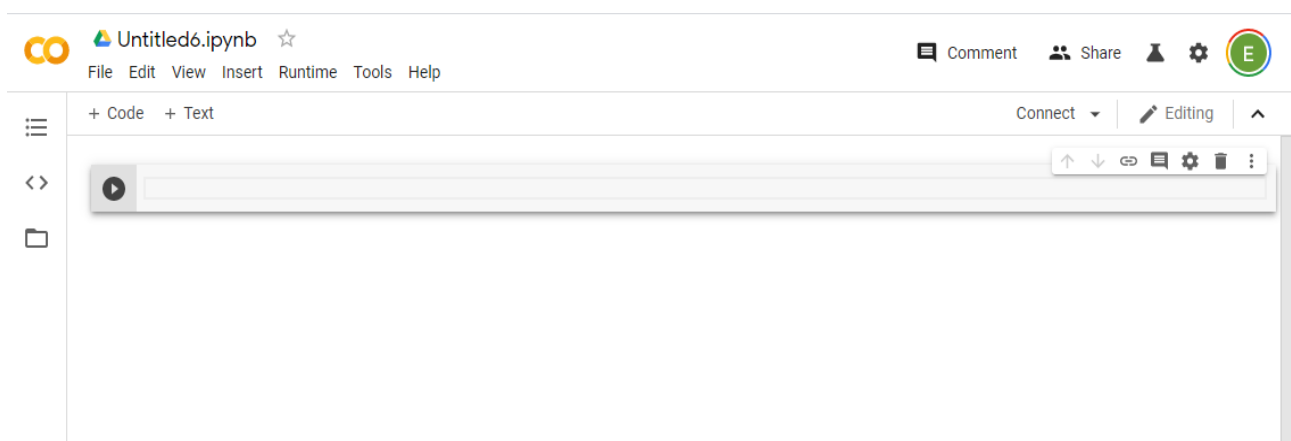


Figura 3. Novo notebook criado ao iniciar uma sessão no Colab.

- Outra forma de entrar no Colab é digitar `colab.research.google.com` no Chrome.
- Essa interface funciona de forma muito parecida com um Notebook Jupyter.

- Ao abrir uma nova sessão no Colab o notebook já vem com uma célula de código.
 - Para inserir novas células de código basta clicar em `+Code` (ver Figura 4).
 - Se quiser inserir uma célula de texto basta clicar em `+Text` (ver Figura 4).

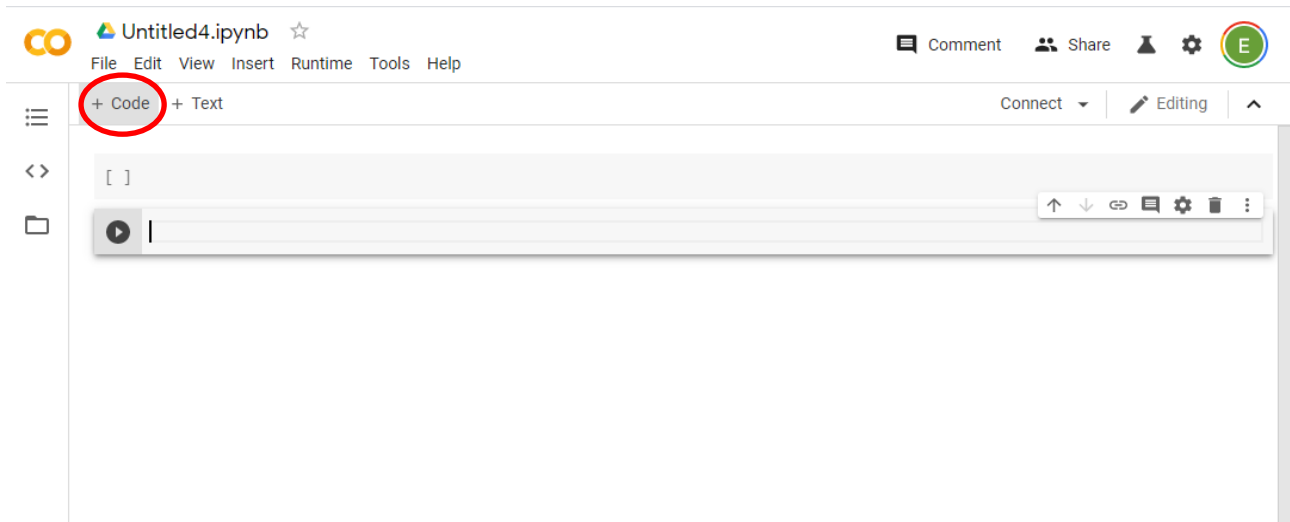


Figura 4. Inserção de células de código e de texto.

- Para ativar a aceleração por GPU deve-se fazer como mostrado na Figura 5.
 - Acessar o menu `Edit > Notebook settings`.
 - Em `Hardware accelerator` mudar para GPU.
 - Neste menu você também pode mudar de Python 3 para Python 2 se quiser.

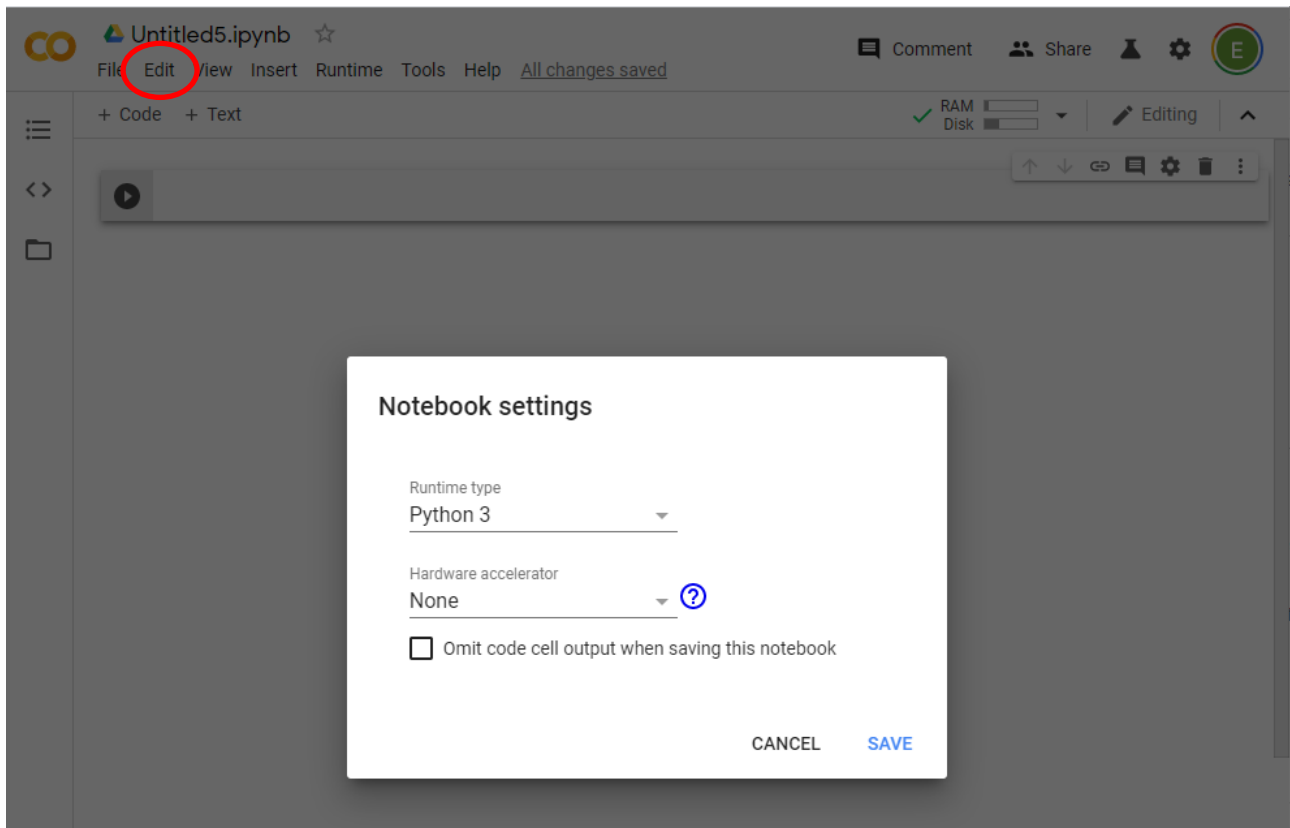


Figura 5. Ativação da GPU.

- Para executar uma célula certifique-se de que o notebook esteja conectado, ou seja, no canto superior esquerdo deve aparecer uma marca verde junto com informações de uso de memória RAM e de disco (ver Figura 6).
- Existem algumas opções para execução de uma célula ou de todo o notebook.
 - Na opção do menu `Runtime` existem várias opções de execução do notebook (Figura 6).
 - Para executar a célula atual, pressionar `SHIFT + ENTER`, ou simplesmente clicar na seta preta que aparece no lado esquerdo da célula, como mostrado na Figura 6.
- Para verificar se está mesmo usando GPU, pode-se fazer um teste usando os comandos: `import tensorflow as tf` e `tf.test.gpu_device_name()`. Escreva esses comandos em uma célula de código e execute, conforme mostrado na Figura 6.

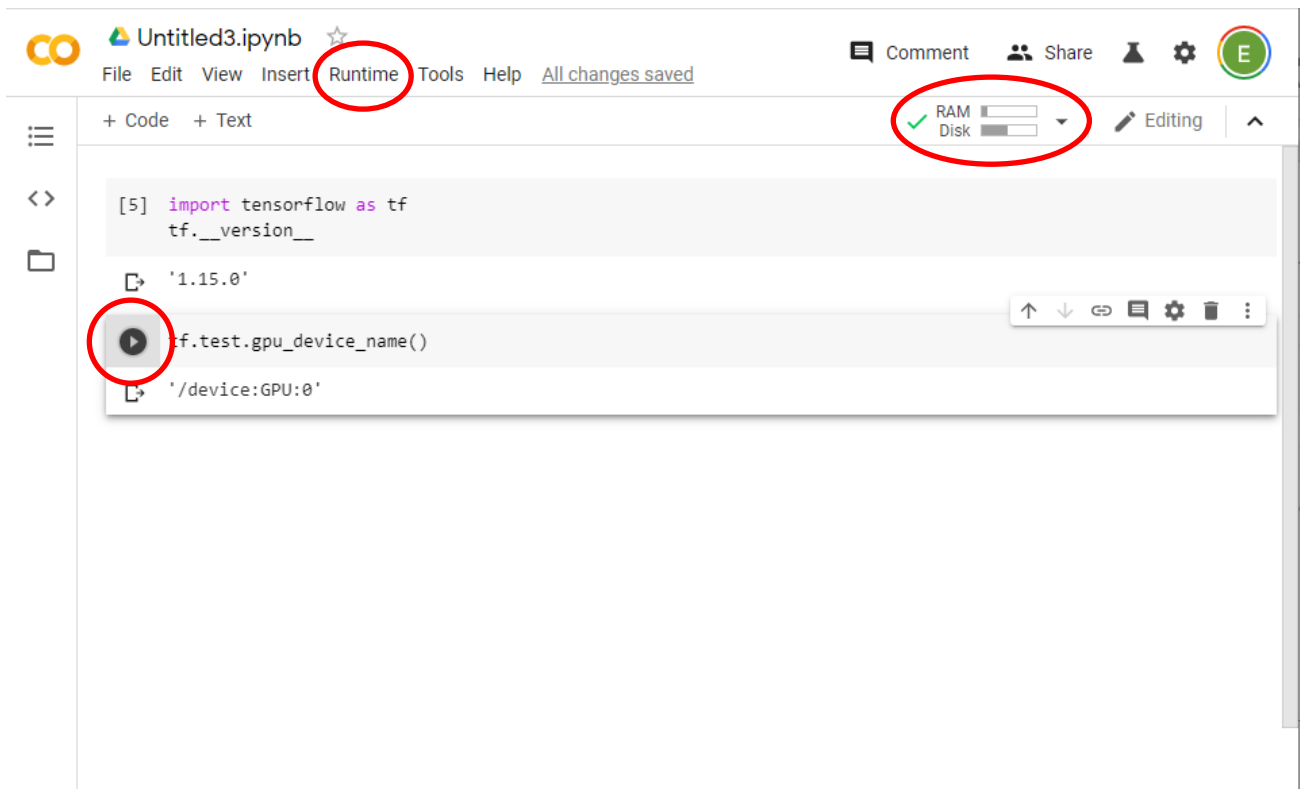


Figura 6. Execução de célula do notebook.

Terminar sessão do Colab

- Para terminar uma sessão do Colab deve-se executar uma célula do notebook com os seguintes comandos:

```
Import os, signal
os.kill(os.getpid(), signal.SIGKILL)
```

- Para que terminar uma sessão do Colab?

É interessante terminar uma sessão do Colab antes de abrir um novo notebook, senão pode ocorrer problemas de memória e também porque existe um número máximo de notebooks que podem estar abertos ao mesmo tempo no Colab.

4. Criar, salvar e abrir notebooks

- Operações com arquivos de notebooks são realizadas na opção `File` do menu ⇒ diversas operações são possíveis. Algumas estão descritas a seguir.
- Para criar um novo notebook clique em `File > New Python 3 notebook`, conforme mostrado na Figura 7.

- Para abrir um notebook existente que está no Google Drive clique em **File > Open notebook...**, conforme mostrado na Figura 7. O notebook deve estar no diretório Colab Notebooks do Google Drive.
- Ao fazer isso aparece a janela mostrada na Figura 8 ⇒ por meio das opções oferecidas nessa janela você pode abrir um notebook trabalhado recentemente, abrir notebooks que estão no diretório Colab Notebooks do seu Google Drive e realizar upload de notebooks do seu computador.

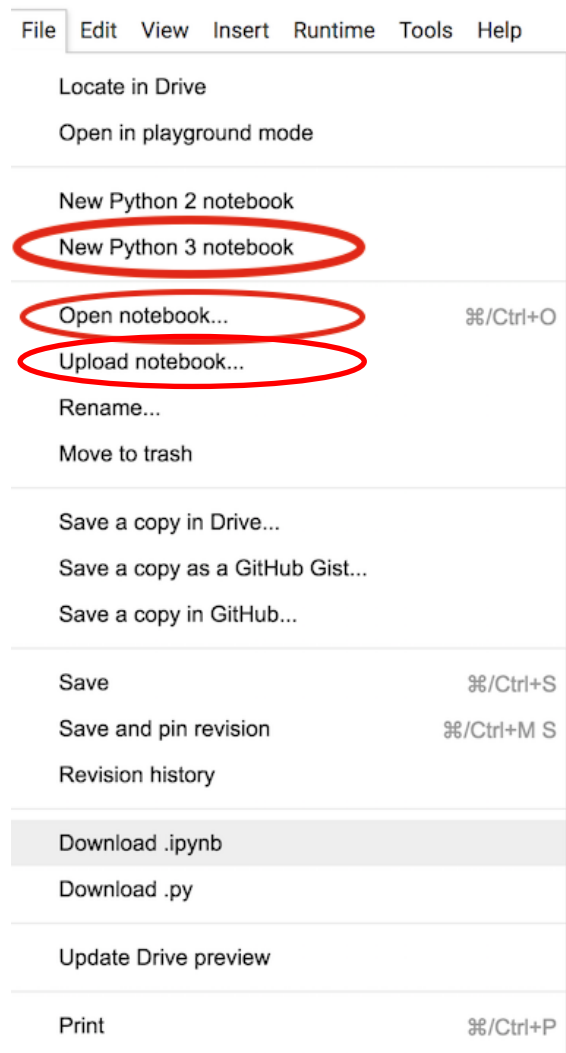


Figura 7. Menu de operações de arquivos de notebooks.

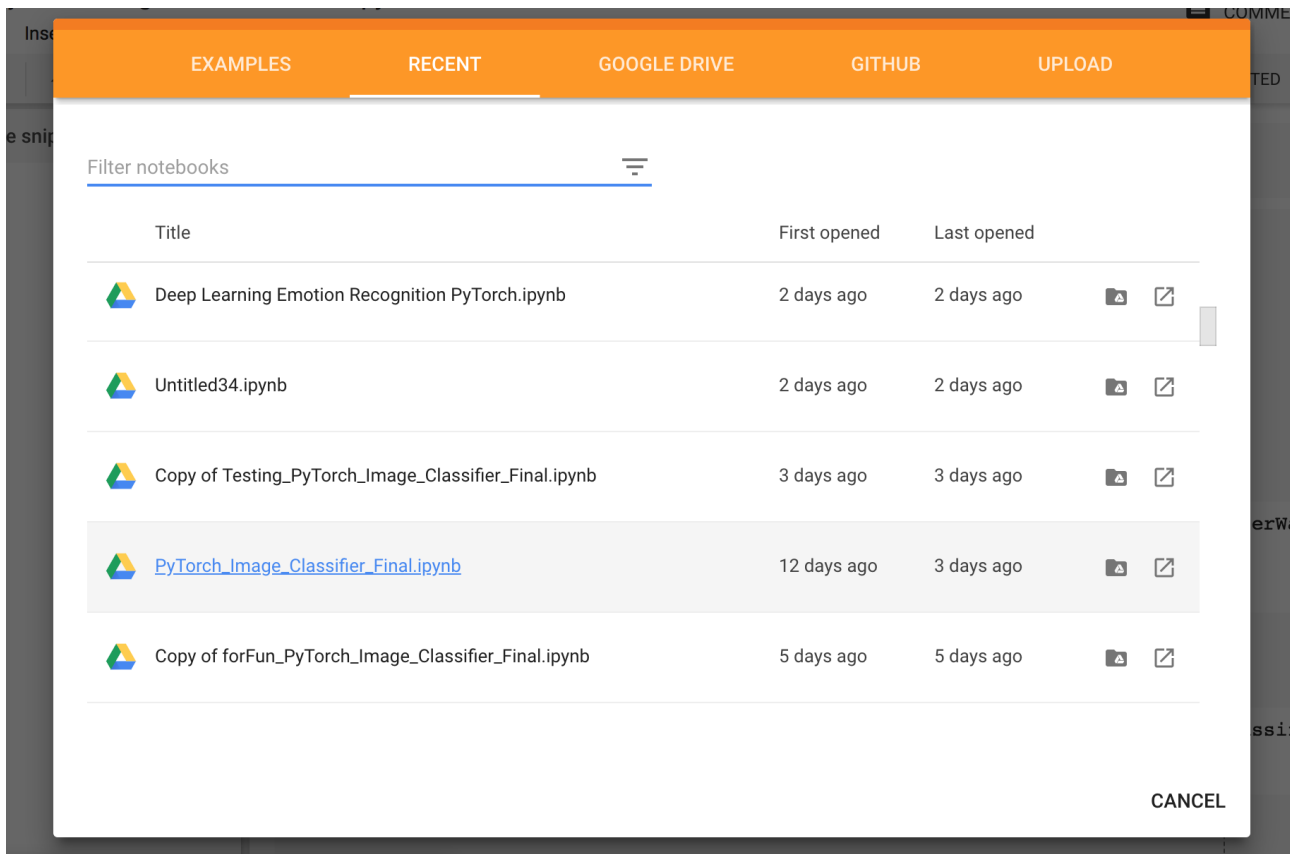


Figura 8. Opções de iniciar notebooks.

- Para abrir um notebook existente que está no seu computador clique em `File > Upload notebook...`. Ao clicar aparece uma nova janela para você navegar e escolher o arquivo, conforme mostrado na Figura 9.

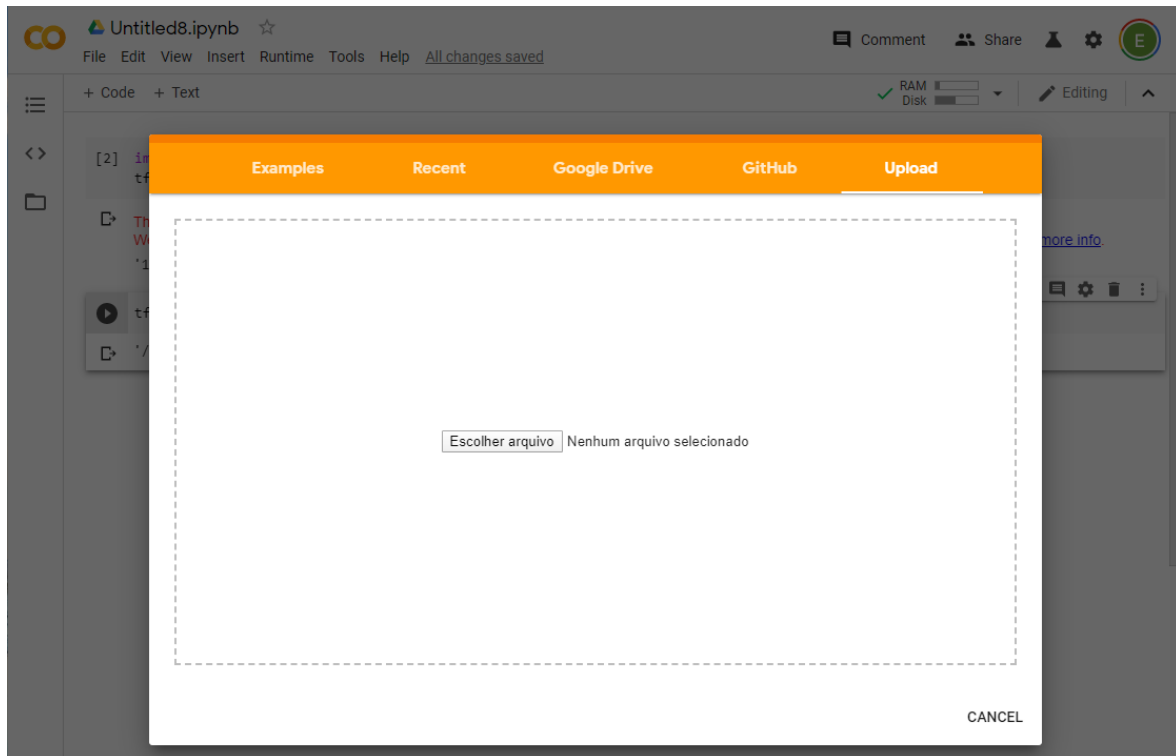


Figura 9. Janela para escolher notebook do seu computador.

- Para salvar um notebook no seu Google Drive clique em `File > Save a copy in Drive...`, conforme mostrado na Figura 10 ⇒ nesse caso o notebook será salvo no diretório Colab Notebooks do Google Drive. Caso esse diretório não exista ele será criado.
- Para salvar um notebook no seu computador clique em `File > Download .ipynb` (ver Figura 10) ⇒ nesse caso o notebook será salvo no diretório de Downloads do seu computador.

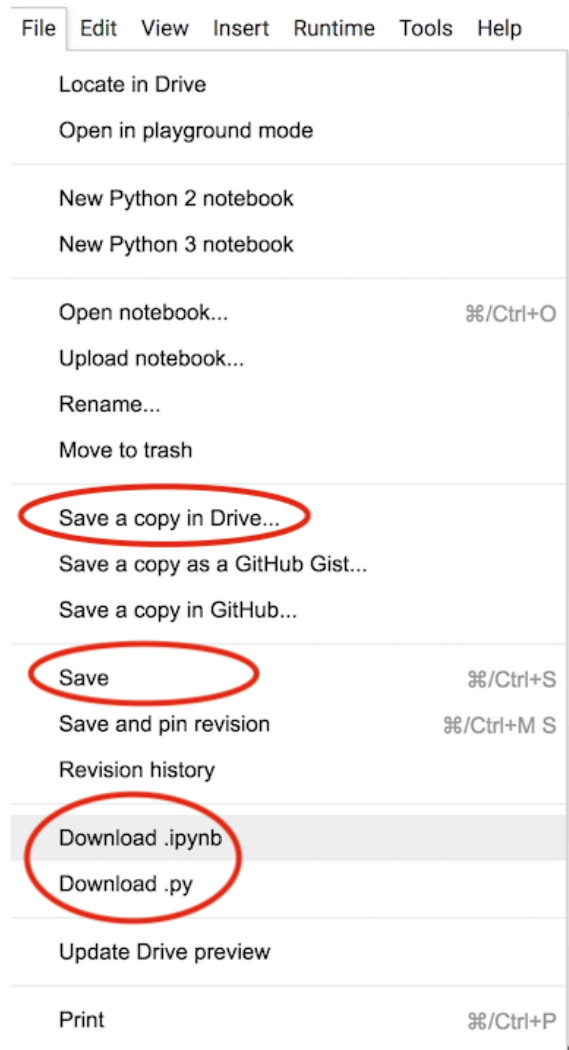


Figura 10. Opções do menu para salvar notebooks.

5. Arquivos

- Em muitos casos são necessários arquivos de dados e scripts Python (código em Python) para executar um notebook ⇒ existe duas formas de acessar arquivos no Colab.
 - Fazendo upload dos arquivos que estão em seu computador.
 - Acessando os arquivos no Google Drive.

Upload de arquivos

- Para fazer upload de arquivos deve-se fazer o seguinte:
 - Clicar no ícone de arquivos localizado no lado esquerdo do seu notebook, como mostrado na Figura 11;

- Clicar na opção Upload ⇒ ao fazer isso abre-se uma janela para escolher os arquivos, navegue no seu computador para selecionar os arquivos.

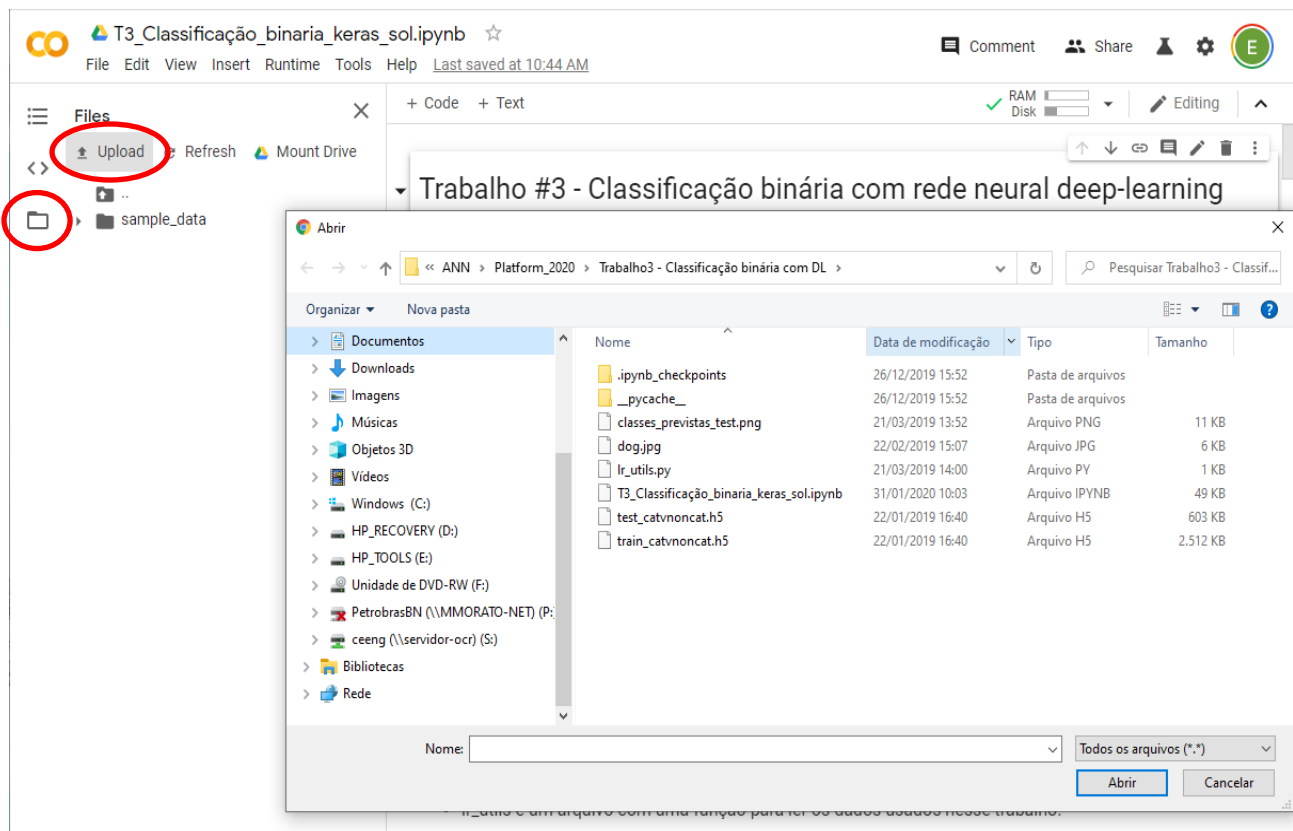


Figura 11. Upload de arquivos.

- Essa forma de usar arquivos é a mais fácil, mas tem alguns inconvenientes:
 - Se sair do Colab ou se fizer alguma alteração no “Notebook settings” esses arquivos são perdidos e você tem que fazer o upload de novo;
 - Se os arquivos de dados forem grandes pode demorar muito para fazer o download.

Acesso de arquivos no Google Drive

- A forma mais eficiente de trabalhar com arquivos no Colab é coloca-los no seu Google Drive.
- Para usar arquivos que estão no seu Google Drive, primeiramente deve-se “montar” o seu Drive. Para isso deve-se fazer o seguinte:
 - No menu do lado esquerdo, clicar na opção Mount Drive como mostrado na Figura 12;
 - Ao fazer isso é inserida uma célula no seu notebook com os comandos abaixo. Essa célula é inserida abaixo da célula onde está o cursor.

```
from google.colab import drive
drive.mount('/content/gdrive')
```

- Ao executar essa célula é gerado um link da internet.
- Clicar no link e seguir as instruções até obter uma chave. Copie a chave na célula do notebook no espaço reservado e aperte a tecla enter.
- Após isso o seu Google Drive fica disponível dentro do diretório drive com nome My Drive (pode ser necessário clicar em Refresh).

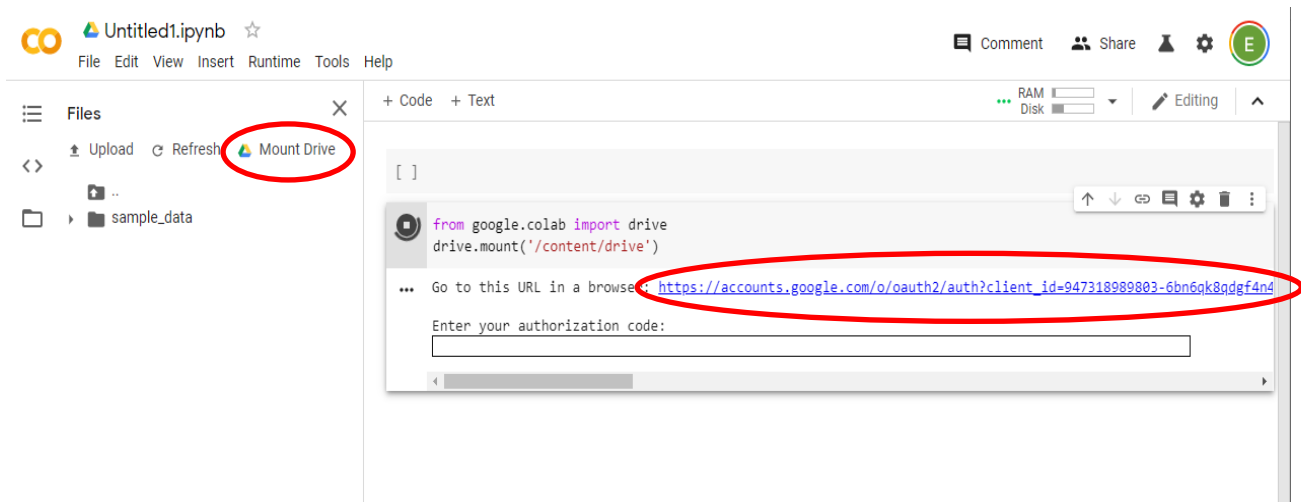


Figura 12. “Montar” o seu Google Drive no Colab.

- O conteúdo do seu Google Drive pode ser listado fazendo:

```
!ls /content/drive/My\ Drive/
```

Note a barra invertida após a palavra My.

- Após montar o seu Google Drive, deve-se alterar o diretório de trabalho fazendo:

```
%cd /content/drive/My Drive/nome_do_diretorio_do_trabalho
```

A operação de alterar o diretório é necessária para poder acessar os seus arquivos pelo seu notebook.

- Para listar o conteúdo do seu diretório de trabalho, após alterar o diretório faz-se:

```
!ls
```

6. Comandos úteis

- No lado esquerdo do Colab existe uma opção que fornece dicas de códigos. Basta clicar no símbolo < > e escrever o que deseja fazer no espaço Filter code snippets, como mostrado na Figura 13.

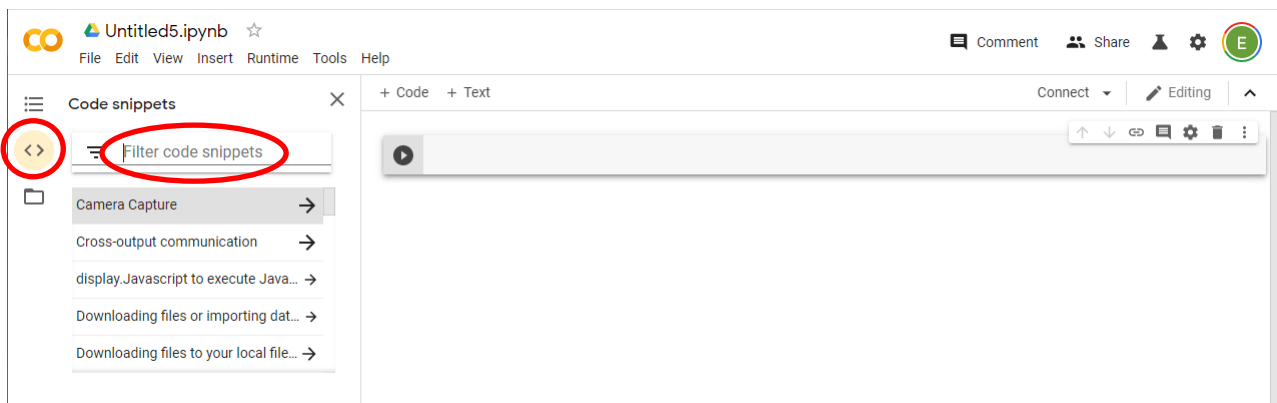


Figura 13. Dicas de códigos.

- Como o TensorFlow é um software aberto não existe muita preocupação do desenvolvedor (Google) de manter compatibilidade entre as versões. Assim, se você estiver usando uma versão do TensorFlow incompatível com o seu programa, você pode impor o uso de uma versão mais antiga.

- Para importar a versão 2 do TensorFlow, mas se ela não estiver acessível importar a versão instalada:

```
try:
    import tensorflow.compat.v2 as tf
except Exception:
    import tensorflow as tf
print(tf.__version__)
```

- O último comando mostra a versão do TensorFlow importada.
- Por exemplo, no caso da versão instalada do TensorFlow for a 2 e o seu programa foi feito para a versão 1, então pode habilitar a versão 2 com o comportamento da versão 1. O código abaixo realiza essa operação:

```
Import tensorflow as tf
print(tf.__version__)
tf.enable_v1_behavior()
```

- Muitas vezes é interessante incluir imagens em células de texto. Para visualizar uma imagem de um arquivo em uma célula de texto faz-se o seguinte:

- Na célula de texto, clicar no ícone de inserir imagem, como mostrado na Figura 14;
- Ao fazer isso é inserido um comando na célula de texto, conforme mostrado na Figura 14;
- Ir no local do seu Google Drive onde se encontra o arquivo da imagem. Clicar no arquivo da imagem com o botão direito do mouse e selecionar a opção Gerar link compartilhável (ver Figura 15);
- Copiar o link na célula de texto no lugar de `https://`;
- Trocar no nome do link `open` por `uc`.

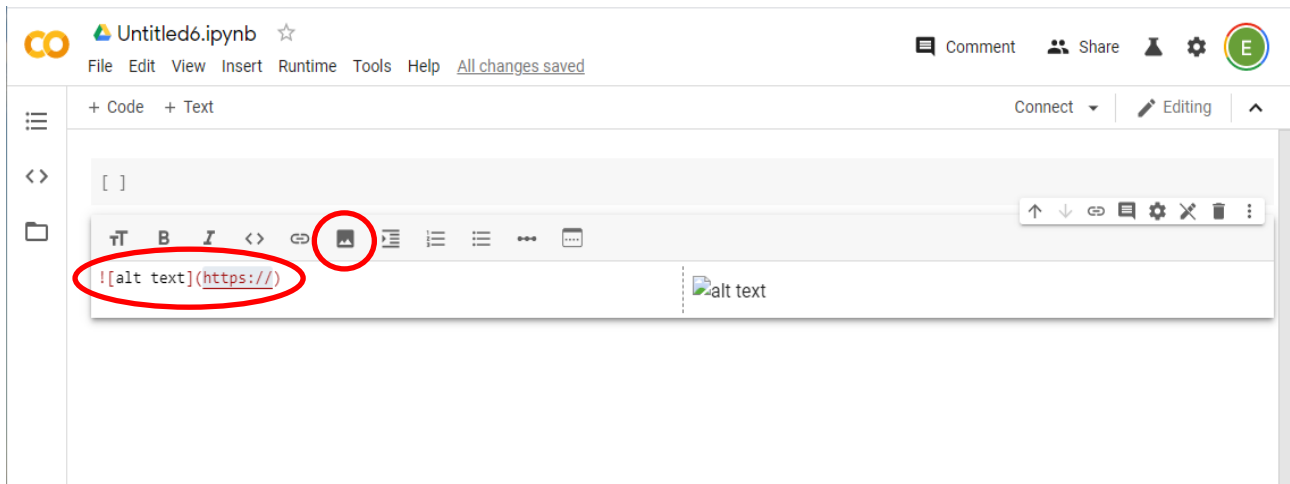


Figura 14. Inserção de imagem em uma célula de texto.

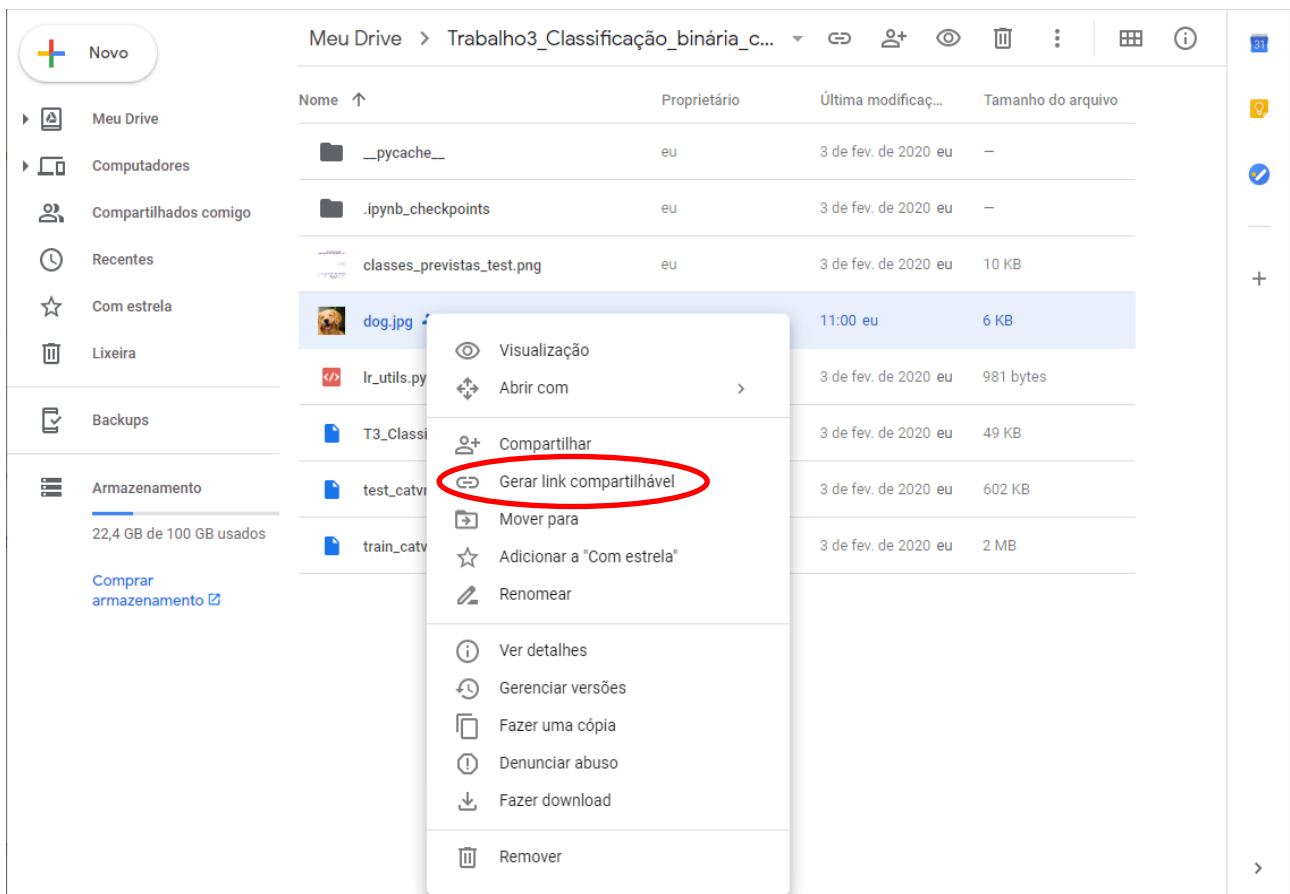


Figura 15. Gerar link compartilhável de um arquivo no Google Drive.

➤ Para criar novos diretórios:

- Criar um diretório dentro do Colab. Note que esse diretório será apagado quando sair do Colab.

```
!mkdir nome_diretorio
```

- Criar um diretório dentro do seu Google Drive.

```
!mkdir /content/drive/My\ Drive/nome_diretorio
```

- Para listar diretórios do Colab:

```
!ls ./sample_data
```

Note o ponto ante do nome.

- Para obter informações sobre memória disponibilizada pelo Colab:

```
!cat /proc/meminfo
```

- Para obter informações sobre a CPU disponibilizada pelo Colab:

```
!cat /proc/cpuinfo
```

- Para verificar as especificações da GPU disponibilizada pelo Colab:

```
from tensorflow.python.client import device_lib  
device_lib.list_local_devices()
```

- Embora a maioria das bibliotecas Python comumente usadas já estão pré-instaladas no Colab, novas bibliotecas podem ser instaladas usando o comando abaixo:

```
!pip install nome_da_biblioteca
```

- Para clonar um repositório do GitHub:

```
!git clone https://github.com/iloonet/tf-pose-estimation.git
```

- Para executar um código Python de um script. O script deve estar no seu Google Drive e o seu Drive deve ter sido montado.

```
!python3 /content/drive/nome_diretorio/script.py
```