

Bootcamp: Analista de Machine Learning**Trabalho Prático****Módulo 1****Fundamentos****Objetivos**

Exercitar os seguintes conceitos trabalhados no Módulo 1:

- ✓ Conhecimento do dataset.
- ✓ Limpeza dos dados.
- ✓ Identificação de Outliers.
- ✓ Análise de regressão linear.

Enunciado

Para o desenvolvimento de qualquer aplicação que utilize os algoritmos de Machine Learning para a solução de problemas reais, são necessárias 7 etapas básicas:

1. Coleta dos dados.
2. Preparação dos dados.
3. Seleção do modelo.
4. Treinamento do modelo.
5. Avaliação do modelo.
6. Sintonia dos parâmetros.
7. Previsão.

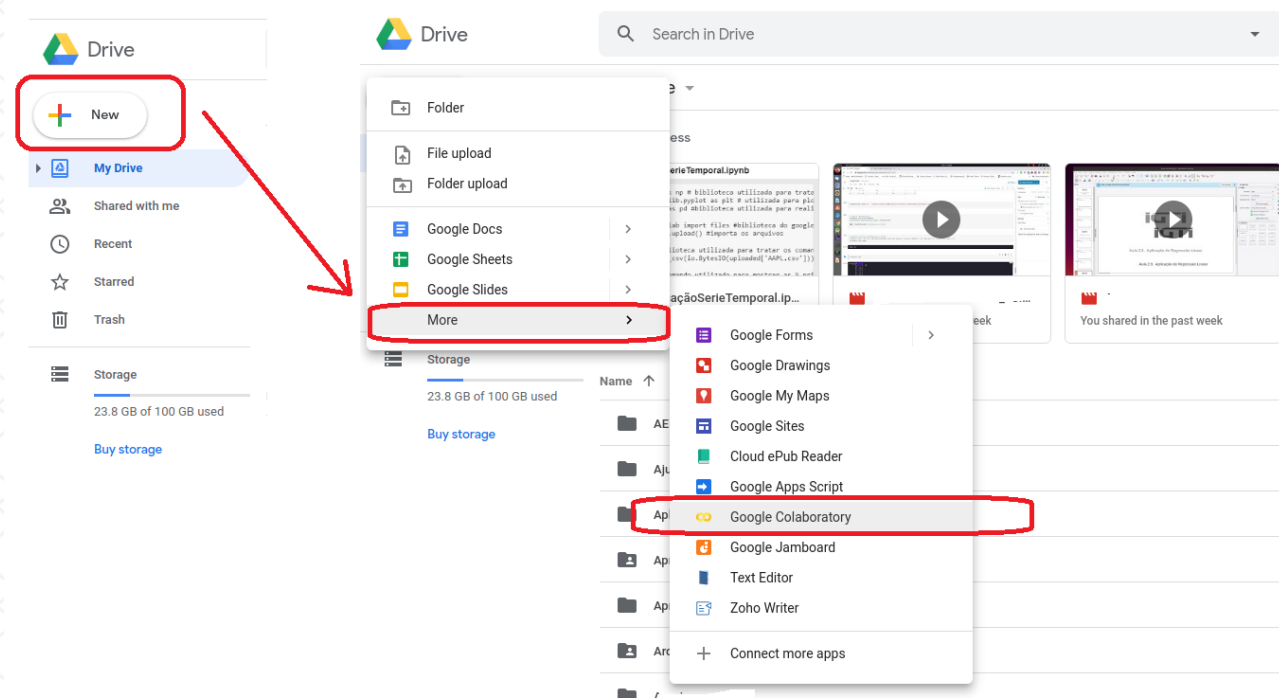
Dentre todas essas etapas, a que demanda um maior esforço por parte do analista/engenheiro de machine learning provavelmente é a etapa de preparação dos dados. Isso ocorre porque é por meio da realização dessa etapa que é possível conhecer o problema em que será empregado o aprendizado de máquina, assim o “dataset” é preparado para um “melhor treinamento” do algoritmo.

Atividades

Os alunos deverão desempenhar as seguintes atividades:

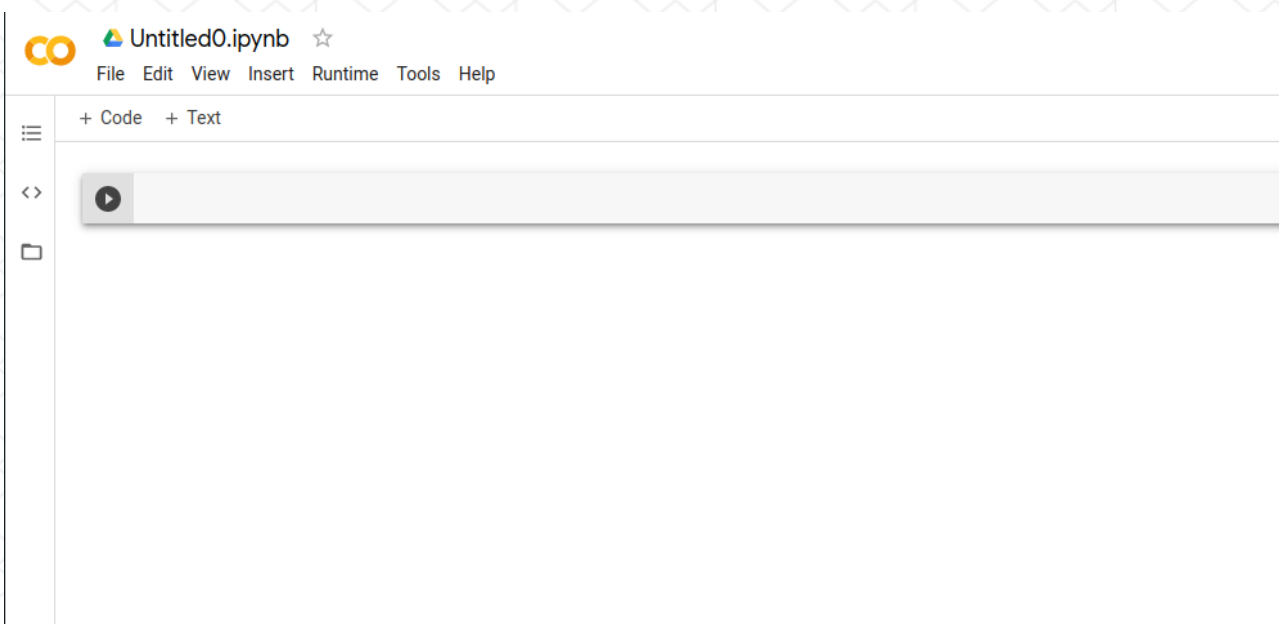
1. Criar uma conta no Google.
2. Acessar o “Google Colaboratory”.
3. Desenvolver o Trabalho Prático.
 - Para este Trabalho Prático, será utilizado o ambiente de desenvolvimento do Google Colab.
 - Para acessar esse ambiente, basta ter uma conta do Google ativa e acessar o Google Drive.
 - Dentro do Google Drive, clique em “**New**”, depois em “**More**” e em seguida selecione “**Google Colaboratory**”. A Figura 1 mostra as etapas necessárias.

Figura 1 – Criando um arquivo no google Colab.



- Após acessar o “Google Colaboratory”, você será direcionado (a) para o ambiente de desenvolvimento do Google. A Figura 2 apresenta a página que deve aparecer ao acessar o ambiente.

Figura 2 – Ambiente do Google Colab.

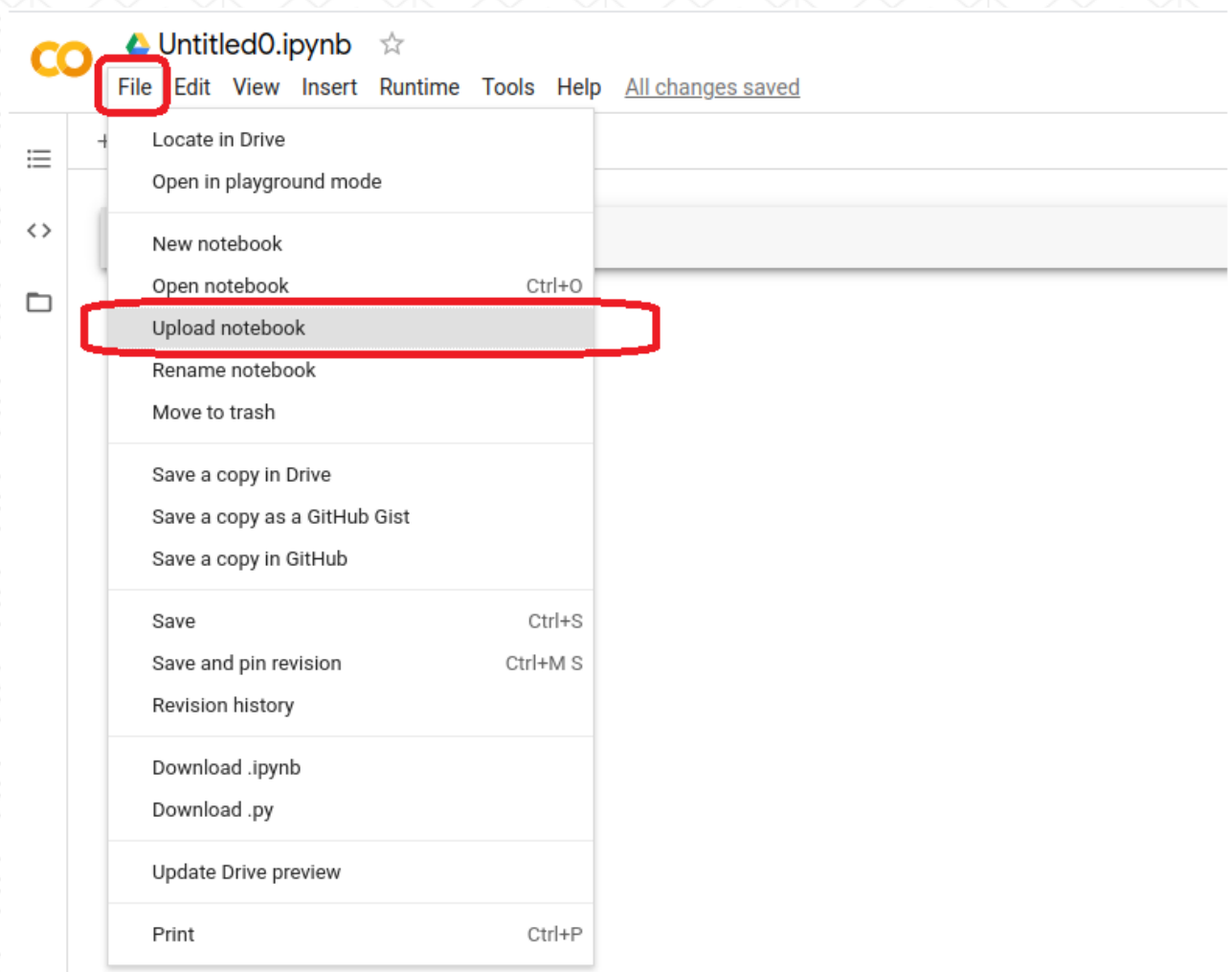


Para essa prática será utilizado o dataset data.csv. Para baixar esse dataset acesse o link abaixo e realize o download dos arquivos data.csv e trabalho_pratico_FAM.ipynb.

https://drive.google.com/open?id=1P0-soA1jXBwSaB6n-ih1CwEC_CjXkPKm

- Com todo o ambiente preparado, é necessário realizar o “**upload**” do arquivo **trabalho_pratico_FAM.ipynb** para o “Google Colab”. Para isso, acesse no canto superior esquerdo o menu “**File**” e clique em “**Upload Notebook**”. No local onde realizou o download dos arquivos anteriores, selecione o arquivo “**trabalho_pratico_FAM.ipynb**”. A Figura 3 mostra como realizar esse procedimento. Após essa etapa, já é possível iniciar o seu Trabalho Prático.

Figura 3 – Upload do arquivo “trabalho_pratico_FAM.ipynb”.



- Para a realização do trabalho, é necessário executar, em sequência, cada uma das células presentes no “**Google Colab**”. Para executar uma célula, selecione a célula desejada e clique o ícone “**play**” (▶) ou pressione “**Ctrl+Enter**”.

Após executar a **célula 3**, será necessário realizar o upload do dataset utilizado para essa prática. Clique no botão “**Escolher Arquivo**” e selecione o arquivo “**data.csv**”. Prosiga executando cada uma das células.

4. Responder as perguntas.

Respostas Finais

Os alunos deverão desenvolver a prática e, depois, responder às seguintes questões objetivas: