kaggle_api_demo

July 11, 2024

1 Kaggle API

At the end of this video, you will be able to: 1. Browse Kaggle datasets through your Jupyter notebook / conda prompt 2. Download Kaggle Datasets directly to your destination path

1.1 Notebook Execution Steps:

1. Install Kaggle_API (ignore if already installed)
2. Set your Kaggle API Token
3. Search for Dataset
4. Download Dataset

2 Step 1: Install Kaggle_API (ignore if already installed)

!pip install kaggle

[1]:
```
!pip install kaggle
```

```
Requirement already satisfied: kaggle in c:\users\shoun\anaconda3\lib\site-
packages (1.5.12)
Requirement already satisfied: tqdm in c:\users\shoun\anaconda3\lib\site-
packages (from kaggle) (4.64.0)
Requirement already satisfied: python-slugify in
c:\users\shoun\anaconda3\lib\site-packages (from kaggle) (5.0.2)
Requirement already satisfied: requests in c:\users\shoun\anaconda3\lib\site-
packages (from kaggle) (2.27.1)
Requirement already satisfied: certifi in c:\users\shoun\anaconda3\lib\site-
packages (from kaggle) (2021.10.8)
Requirement already satisfied: python-dateutil in
c:\users\shoun\anaconda3\lib\site-packages (from kaggle) (2.8.2)
Requirement already satisfied: urllib3 in c:\users\shoun\anaconda3\lib\site-
packages (from kaggle) (1.26.9)
Requirement already satisfied: six>=1.10 in c:\users\shoun\anaconda3\lib\site-
packages (from kaggle) (1.16.0)
Requirement already satisfied: text-unidecode>=1.3 in
c:\users\shoun\anaconda3\lib\site-packages (from python-slugify->kaggle) (1.3)
Requirement already satisfied: idna<4,>=2.5 in
c:\users\shoun\anaconda3\lib\site-packages (from requests->kaggle) (3.3)
Requirement already satisfied: charset-normalizer~=2.0.0 in
```

```
c:\users\shoun\anaconda3\lib\site-packages (from requests->kaggle) (2.0.4)
Requirement already satisfied: colorama in c:\users\shoun\anaconda3\lib\site-
packages (from tqdm->kaggle) (0.4.4)
```

3 Step 2: Set your Kaggle API Token

- Download `kaggle.json` file from Kaggle
- Create `.kaggle` folder in ROOT directory
- Copy-Paste `kaggle.json` to ROOT directory

4 Step 3: Search Dataset

See List of Available Datasets !kaggle datasets list -s "dataset name"

[2]:
```
!kaggle datasets list -s "phone"
```

```
ref                                                         title
size  lastUpdated          downloadCount  voteCount  usabilityRating
----------------------------------------------------------
----------------------------------------------  -----  -------------------
-------------  ---------  ---------------
grikomsn/amazon-cell-phones-reviews                         Amazon Cell Phones
Reviews                             9MB  2019-12-26 22:21:16          13296
199  1.0
PromptCloudHQ/amazon-reviews-unlocked-mobile-phones     Amazon Reviews: Unlocked
Mobile Phones                33MB  2017-01-11 10:22:30          11903        162
1.0
marcodena/mobile-phone-activity                             Mobile phone activity in
a city                      242MB  2019-11-14 06:11:32          11746        245
0.7058824
prasertk/mobile-phone-rating                                Mobile phone rating
3KB  2022-02-13 08:02:17            764         18  1.0
arwinneil/gsmarena-phone-dataset                            GSMArena Phone Dataset
635KB  2017-06-29 07:09:31           5287         71  0.85294116
mfekadu/darpa-timit-acousticphonetic-continuous-speech  DARPA TIMIT Acoustic-
Phonetic Continuous Speech     829MB  2019-06-05 02:01:06           5866
116  0.8235294
msainani/gsmarena-mobile-devices                            GSMArena Mobile Phone
Devices                           1MB  2020-06-24 15:12:02           2075
49  0.9411765
muhammedtausif/best-selling-mobile-phones                   Best Selling Mobile
Phones                              2KB  2022-05-22 17:31:44           1209
43  1.0
ruchi798/mobile-phone-tweets                                Mobile Phone Tweets
4MB  2021-12-15 14:15:02            405         30  1.0
uciml/human-activity-recognition-with-smartphones           Human Activity
Recognition with Smartphones          25MB  2019-11-13 20:04:13          31851
550  0.7058824
```

```
artempozdniakov/ukrainian-market-mobile-phones-data     Mobile Phones Data
31KB  2021-02-05 10:48:09           2202         27  1.0
malekzadeh/motionsense-dataset                          MotionSense Dataset :
Smartphone Sensor Data - HAR   72MB  2018-03-15 20:50:21           8658
225  0.7647059
frtgnn/rural-residents-daily-mobile-phone-data          Rural Residents Daily
Mobile Phone Data             345KB  2020-04-11 22:23:26            189
63  1.0
ak47bluestack/amazonphonedataset                        Amazon-Phone-Dataset
4MB  2019-07-16 07:56:15           1565         24  0.88235295
khaiid/android-phones                                   Android Phones
11KB  2022-01-03 03:35:31            259         27  1.0
devsubhash/flipkart-mobiles-dataset                     Flipkart Mobiles Dataset
54KB  2022-03-20 10:46:28           1432         35  1.0
iabhishekofficial/mobile-price-classification           Mobile Price
Classification                           71KB  2018-01-28 08:44:24          92141
1246  0.7058824
google/android-smartphones-high-accuracy-datasets       Android smartphones high
accuracy GNSS datasets       1GB  2020-12-23 01:51:11           2245        179
0.875
nkitgupta/flipkart-cell-phone-reviews                   Flipkart Cell Phone
Reviews                          3MB  2022-07-05 06:59:34             41
18  1.0
ruthgn/bank-marketing-data-set                          Bank Marketing Data Set
282KB  2021-10-20 00:33:20           1847         50  1.0
```

5 Step 4: Download the Dataset

!kaggle datasets download -d "kaggle_dataset_name"

[3]:
```
!kaggle datasets download -d "prasertk/mobile-phone-rating"
```

```
Downloading mobile-phone-rating.zip to
C:\Users\shoun\Jupyter\Kaggle\plain_data\kaggle_api_demo


  0%|          | 0.00/3.35k [00:00<?, ?B/s]
100%|##########| 3.35k/3.35k [00:00<00:00, 1.14MB/s]
```

[ ]:
```

```

6 Unzip and Read the Data

Unzip the ".zip" file using `zipfile` library example: `with zipfile.ZipFile("source_file_path.zip", "r") as file: file.extractall("destination_path")`

[4]:
```
import zipfile
```

[5]:
```
with zipfile.ZipFile("mobile-phone-rating.zip", "r") as file:
    file.extractall("phones")
```

[ ]:
```
# See the Contents of the Extracted File
# import os
# os.listdir(destination_path)
```

[7]:
```
import os

os.listdir("phones")
```

[7]:
```
['mobile phone rating by dxo.csv']
```

[ ]:
```
# Read Dataset
```

[8]:
```
import pandas as pd
```

[9]:
```
pd.read_csv("phones/mobile phone rating by dxo.csv")
```

[9]:

	model	price	launch	camera	selfie	audio	display	battery
0	Huawei P50 Pro	$907	Jul 2021	144.0	106.0	NaN	93.0	NaN
1	Xiaomi Mi 11 Ultra	$1200	Mar 2021	143.0	94.0	71.0	87.0	69.0
2	Huawei Mate 40 Pro+	$1363	Oct 2020	139.0	NaN	NaN	NaN	NaN
3	Apple iPhone 13 Pro Max	$1099	Sep 2021	137.0	99.0	75.0	99.0	89.0
4	Apple iPhone 13 Pro	$999	Sep 2021	137.0	99.0	75.0	98.0	76.0
..	…	…	…	…	…	…	…	…
218	Huawei P40 Lite	$279	Feb 2020	NaN	NaN	NaN	NaN	73.0
219	Wiko Power U20	$159	Jan 2021	NaN	NaN	NaN	NaN	70.0
220	Motorola Moto G9 Power	$229	Nov 2020	NaN	NaN	NaN	NaN	70.0
221	Realme C11	$119	Jun 2020	NaN	NaN	NaN	NaN	61.0
222	Xiaomi Redmi 9	$159	Jun 2020	NaN	NaN	NaN	NaN	54.0

[223 rows x 8 columns]