
kaggledatasets

Dec 18, 2019

1	Installation	1
2	Configuration	3
3	Frequently Asked Questions	5
4	Contributing Guide	7
5	Persons of Interest	9
6	kaggledatasets	11
7	kaggledatasets.core	13
8	kaggledatasets.core.dataset	15
9	kaggledatasets.core.config	17
10	kaggledatasets.core.downloader	19
11	kaggledatasets.core.fileops	21
12	kaggledatasets.structured	23
13	kaggledatasets.image	25
14	kaggledatasets.audio	27
15	kaggledatasets.text	29
16	Indices and tables	31

CHAPTER 1

Installation

kaggledatasets requires Python 3.5+

kaggledatasets is available in the Python Package Index via

```
$ pip install kaggledatasets
```

The easiest way to get started on most systems is to create a *virtualenv*

```
$ python3 -m venv kd_venv
$ source kd_venv/bin/activate
(kd_venv) $ pip install kaggledatasets
```

This will install a version of all TF and PyTorch dependencies depending on your system. See [kaggledatasets](#) for more information.

If you need a different version of *kaggledatasets*, follow the instructions on the [kaggledatasets website](#) to install the appropriate version of [kaggledatasets.github.io](#)

1.1 Install From Source

```
$ git clone git@github.com:kaggledatasets/kaggledatasets.git
$ cd kaggledatasets
$ source kd_venv
(kd_venv) $ python3 setup.py install
```

Once that is installed, you can run the unit tests. We recommend using X as a runner.

To resume development in an already checked-out repo:

```
$ cd kaggledatasets
$ source kd_venv
```

To exit the virtual environment:

```
(kd_venv) $ deactivate
```

Coming Soon

1.2 Cloud VM Setup

This guide will cover all the setup work you have to do in order to be able to easily install kaggledatasets on a cloud VM . *Note that while these instructions worked when they were written, they may become incorrect or out of date. If they do, please send us a Pull Request!*

After following these instructions, you should be good to either follow the [Installation](#) instructions or the [Install From Source](#) instructions

1.2.1 Amazon Web Services

Coming Soon

1.2.2 Google Cloud Engine

Coming Soon

1.2.3 Microsoft Azure

Coming Soon

2.1 Local System (Windows/Linux/macOS)

This guide will cover all the setup work you have to do in order to be able to easily install and configure kaggle datasets on your local machine.

2.2 Google Colab Setup

This guide will cover all the setup work you have to do in order to be able to easily install and configure kaggle datasets on Google Colab

CHAPTER 3

Frequently Asked Questions

First off, thanks for taking the time to contribute!

4.1 Before you start

- Comment on the issue that you plan to work on so we can assign it to you and there isn't unnecessary duplication of work.
- When you plan to work on something larger (for example, adding new features to Dataset Class), please respond on the issue (or create one if there isn't one) to explain your plan and give others a chance to discuss.
- If you're fixing some smaller issue - please check the list of pending Pull Requests to avoid unnecessary duplication.

4.2 How can you help

You can help in multiple ways: * Reproducing bugs, finding its root cause and providing fixes to that, this will be appreciated a lot (see the issues with label: *bug*) * Sending Pull Requests for new kaggle datasets and/or requested features (see the issues with label: *dataset request* or *enhancement*) * Doing Code Reviews on the Pull Requests from the developers of this community and verifying if PRs are working correctly or not

4.3 Datasets

Adding a Kaggle Dataset is a great way of making it more accessible to the various communities. **Add a new Kaggle Dataset** <>'_ guide will be available soon.

4.4 Tests

We use [pylint](#) for ensuring kaggledatasets is nice and easy to use and work on long-term, all modules should have clear tests for public members.

4.5 Pull Requests

All contributions are done through Pull Requests here on GitHub.

4.6 Code Reviews

All submissions, including submissions by project members, require review. We use GitHub pull requests for this purpose. Consult [GitHub Help](#) for more information on using pull requests.

5.1 Looking for contributors who can be our Persons of Interest

5.2 General Maintainers

- Omkar Prabhu ([prabhuomkar](#))

5.3 Module-level maintainers

5.3.1 Core

- Omkar Prabhu ([prabhuomkar](#))

5.3.2 Structured

- Omkar Prabhu ([prabhuomkar](#))

5.3.3 Tutorials

CHAPTER 6

kaggledatasets

kaggledatasets

CHAPTER 7

kaggledatasets.com

CHAPTER 8

`kaggledatasets.core.dataset`

CHAPTER 9

`kaggledatasets.core.config`

CHAPTER 10

`kaggledatasets.core.downloader`

CHAPTER 11

kaggledatasets.core.fileops

CHAPTER 12

kaggledatasets.structured

The `kaggledatasets.structured` subpackage consists of popular datasets and common functions for structured datasets like *CSV*, *JSON*, *SQLITE*, etc.

CHAPTER 13

kaggledatasets.image

The `kaggledatasets.image` subpackage consists of popular datasets and common image transformations for image datasets.

CHAPTER 14

kaggledatasets.audio

The `kaggledatasets.audio` subpackage consists of popular datasets and common audio transformations.

CHAPTER 15

kaggledatasets.text

The `kaggledatasets.text` subpackage consists of data processing utilities and popular datasets for natural language.

CHAPTER 16

Indices and tables

- `genindex`
- `modindex`
- `search`