
tutorials001 Documentation

Release latest

Nov 17, 2021

CONTENTS:

1	Introduction	1
2	Installation/Usage:	3
3	Import data-scraper	5
4	Train	7
5	Run	9
	Index	11

INTRODUCTION

With `data-scraper` you can scrape any website without the need of inspecting web elements or parsing HTML using BeautifulSoup etc. With just URLs as input you get JSON data as output. First you need to train the scraper for particular website & then run it.

It uses selenium to automate the things. You can use its inbuilt functions in a very easy way.

INSTALLATION/USAGE:

You can find this package on Pypi (see [here](#)).

Command to install :- `pip install data-scraper`

IMPORT DATA-SCRAPER

```
from data_scraper import *
```


TRAIN

It takes two URLs of 2 similiar pages to train the scraper.

status: True or False

id: id of the scraper

Here is the code:-

```
scraper.train(link1, link2)
```

Parameters

- **link1** (*str*) – First link on which scraper to run
- **link2** (*str*) – Similiar link to link1

Returns {"status": True, "id": "Adfef343JDJSDJ"}

Return type dict

RUN

It runs the scraper and gives data in response.

Here is the code:-

```
scraper.run(link1, id='Adfef343JDJSDJ')
```

Parameters

- **link1** (*str*) – Link of which data to scraper
- **id** (*str*) – id of the scraper got in training

Returns {“data1”:”data1”,... }

Return type dict

INDEX

B

built-in function
 scraper.run(), 9
 scraper.train(), 7

S

scraper.run()
 built-in function, 9
scraper.train()
 built-in function, 7