steppy Documentation

Release 0.1.3

minerva.ml

Jun 15, 2018

Module contents:

1 API documentation 1

2 What is Steppy? 3

3 What problem steppy solves? 5

4 Start using steppy 7
4.1 Installation 7

5 Resources 9

6 Feature Requests 11

7 Roadmap 13

8 Related projects 15

9 Contributing 17

10 Terms of use 19

CHAPTER 1

API documentation

- genindex
- modindex
- search

CHAPTER 2

What is Steppy?

Steppy is a lightweight, open-source, Python 3 library for fast and reproducible experimentation. It lets data scientist focus on data science, not on software development issues. Steppy's minimal interface does not impose constraints, however, enables clean machine learning pipeline design.

CHAPTER 3

What problem steppy solves?

In the course of the project, data scientist faces multiple problems. Difficulties with reproducibility and lack of the ability to prepare experiments quickly are two particular examples. Steppy address both problems by introducing two simple abstractions: `Step` and `Tranformer`. We consider it minimal interface for building machine learning pipelines.

`Step` is a wrapper over the transformer and handles multiple aspects of the execution of the pipeline, such as saving intermediate results (if needed), checkpoiting the model during training and much more. `Tranformer` in turn, is purely computational, data scientist-defined piece that takes an input data and produces some output data. Typical Transformers are neural netowrk, machine learning algorithms and pre- or post-processing routines.

CHAPTER 4

Start using steppy

4.1 Installation

Steppy requires `python3.5` or above.

```
pip3 install steppy
```

(you probably want to install it in your virtualenv*)*

CHAPTER 5

Resources

1. Documentation
2. Source
3. Bugs reports
4. Feature requests
5. Tutorial notebooks (their repository):
    - Getting started
    - Steps with multiple inputs
    - Advanced adapters
    - Caching and persistance
    - Steppy with Keras

CHAPTER 6

Feature Requests

Please send us your ideas on how to improve steppy library! We are looking for your comments here: Feature requests.

CHAPTER 7

Roadmap

At this point steppy is early-stage library heavily tested on multiple machine learning challenges (data-science-bowl, toxic-comment-classification-challenge, mapping-challenge) and educational projects (minerva-advanced-data-scientific-training).

We are developing steppy towards practical tool for data scientists who can run their experiments easily and change their pipelines with just few manipulations in the code.

CHAPTER 8

Related projects

We are also building steppy-toolkit, a collection of high quality implementations of the top deep learning architectures -> all of them with the same, intuitive interface.

CHAPTER 9

Contributing

You are welcome to contribute to the Steppy library. Please check CONTRIBUTING for more information.

CHAPTER 10

Terms of use

Steppy is MIT-licesed.