CS272 Project Documentation

Release 0.1.0

Michael Tao-Yi Lee

Mar 20, 2020

Contents:

CS272 Project

NLP Final Project

- Free software: MIT license

- Documentation: https://cs272-project.readthedocs.io.

1.1 Quick Start

1.1.1 Replication Only

Following the steps below to setup training environment

```
mkdir work_directory
cd work_directory
# create virtual environment under work_directory, naming it to "venv"
python -m venv venv
source venv/bin/activate
# install the package
pip install cs272-project
# write configuration file to a working directory
cs272_project_cli write-config --outfile /home/tylee/PycharmProjects/nlp_workspace
```

1.1.2 Development

First of all, please fork the project if you are interested in extending its functionalities. After that, you may clone the repository with:

```
git clone git@github.com:<user_name>/nlp_project.git
```

Pull requests welcome!

1.1.3 Dependencies

1. PyTorch == 1.4.0
2. YouTokenToMe (yttm) == 1.0.6
3. Huggingface Transformers == 2.4.1
4. Click >= 0.7

1.2 Features

- TODO

1.3 Credits

This package was created with Cookiecutter and the audreyr/cookiecutter-pypackage project template.

Installation

2.1 Stable release

To install CS272 Project, run this command in your terminal:

```
$ pip install cs272_project
```

This is the preferred method to install CS272 Project, as it will always install the most recent stable release.

If you don't have pip installed, this Python installation guide can guide you through the process.

2.2 From sources

The sources for CS272 Project can be downloaded from the Github repo.

You can either clone the public repository:

```
$ git clone git://github.com/taoyilee/cs272_project
```

Or download the tarball:

```
$ curl -OJL https://github.com/taoyilee/cs272_project/tarball/master
```

Once you have a copy of the source, you can install it with:

```
$ python setup.py install
```

Usage

To use CS272 Project in a project:

```python
import cs272_project
```

Contributing

Contributions are welcome, and they are greatly appreciated! Every little bit helps, and credit will always be given.

You can contribute in many ways:

4.1 Types of Contributions

4.1.1 Report Bugs

Report bugs at https://github.com/taoyilee/cs272_project/issues.

If you are reporting a bug, please include:

- Your operating system name and version.
- Any details about your local setup that might be helpful in troubleshooting.
- Detailed steps to reproduce the bug.

4.1.2 Fix Bugs

Look through the GitHub issues for bugs. Anything tagged with "bug" and "help wanted" is open to whoever wants to implement it.

4.1.3 Implement Features

Look through the GitHub issues for features. Anything tagged with "enhancement" and "help wanted" is open to whoever wants to implement it.

4.1.4 Write Documentation

CS272 Project could always use more documentation, whether as part of the official CS272 Project docs, in docstrings, or even on the web in blog posts, articles, and such.

4.1.5 Submit Feedback

The best way to send feedback is to file an issue at https://github.com/taoyilee/cs272_project/issues.

If you are proposing a feature:

- Explain in detail how it would work.

- Keep the scope as narrow as possible, to make it easier to implement.

- Remember that this is a volunteer-driven project, and that contributions are welcome :)

4.2 Get Started!

Ready to contribute? Here's how to set up *cs272_project* for local development.

1. Fork the *cs272_project* repo on GitHub.

2. Clone your fork locally:

```
$ git clone git@github.com:your_name_here/cs272_project.git
```

3. Install your local copy into a virtualenv. Assuming you have virtualenvwrapper installed, this is how you set up your fork for local development:

```
$ mkvirtualenv cs272_project
$ cd cs272_project/
$ python setup.py develop
```

4. Create a branch for local development:

```
$ git checkout -b name-of-your-bugfix-or-feature
```

Now you can make your changes locally.

5. When you're done making changes, check that your changes pass flake8 and the tests, including testing other Python versions with tox:

```
$ flake8 cs272_project tests
$ python setup.py test or pytest
$ tox
```

To get flake8 and tox, just pip install them into your virtualenv.

6. Commit your changes and push your branch to GitHub:

```
$ git add .
$ git commit -m "Your detailed description of your changes."
$ git push origin name-of-your-bugfix-or-feature
```

7. Submit a pull request through the GitHub website.

4.3 Pull Request Guidelines

Before you submit a pull request, check that it meets these guidelines:

1. The pull request should include tests.

2. If the pull request adds functionality, the docs should be updated. Put your new functionality into a function with a docstring, and add the feature to the list in README.rst.

3. The pull request should work for Python 3.5, 3.6, 3.7 and 3.8, and for PyPy. Check https://travis-ci.com/taoyilee/cs272_project/pull_requests and make sure that the tests pass for all supported Python versions.

4.4 Tips

To run a subset of tests:

```
$ pytest tests.test_cs272_project
```

4.5 Deploying

A reminder for the maintainers on how to deploy. Make sure all your changes are committed (including an entry in HISTORY.rst). Then run:

```
$ bump2version patch # possible: major / minor / patch
$ git push
$ git push --tags
```

Travis will then deploy to PyPI if tests pass.

Credits

5.1 Development Lead

- Michael Tao-Yi Lee <taoyil@uci.edu>

5.2 Contributors

None yet. Why not be the first?

CHAPTER 6

History

6.1 0.1.0 (2020-02-15)

- First release on PyPI.

6.2 0.2.0 (2020-03-01)

- Change the fine tuning method to work with GPT2TANDAModel() which is a dual head model for AS2 and ODQA

Indices and tables

- genindex
- modindex
- search