u/prequelmemes$_{\text{bot}}$ Documentation

thecsw

May 11, 2018

Contents

Hello there!

Welcome to the source page of u/prequelmemes$_{bot}$! Nice to meet you. First of all, let me talk you through what the bot does and details later. **If you want** to help the bot, please go down to the **TODO** list and find your way to contribute! *Thanks!*
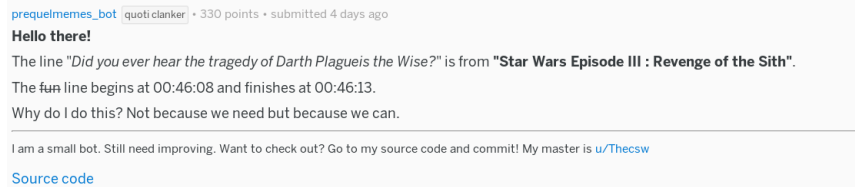
Figure 1: A comment picture.

1 Getting started

These instructions will get you a copy of the project up and running on your local machine for development and testing purposes. See deployment for notes on how to deploy the project on a live system.

1.1 Prerequisites

```
sudo pip install praw
sudo pip install python-opencv
sudo pip install pytesseract
sudo pip install tesseract-ocr
sudo pip install Pillow
sudo pip install tqdm
```

OR

```
pip install --upgrade -r requirements.txt
```

1. praw is Python Reddit API Wrapper. This will be the main and only package to connect to Reddit's API and extract desired data.

2. python-opencv is used for image transformations and computer vision problems.

3. pytesseract is a python wrapper for Google's Tesseract-OCR.

4. Pillow is the Python Imaging Library by Fredrik Lundh and Contributors.

5. tqdm is used for fancy progress bars.

1.2 Other Dependencies

Tesseract engine should be installed on a local machine to run the text recognition properly. We will also install the tesseract OCR trained languages for better accuracy and we will install only the English packages. For more information about other languages, please refer to tesseract's official repository on Github.

1.2.1 Linux

1. Debian, Ubuntu (aptitude)

```
sudo apt-get install tesseract-ocr
sudo apt-get install tesseract-ocr-eng
```

2. Arch Linux (pacman)

```
sudo pacman -S tesseract
sudo pacman -S tesseract-data-eng
```

1.2.2 Mac OS (homebrew)

```
brew install tesseract
```

2 Installing

The only thing that needs to be done before execution is the config. The only thing that needs to be done before execution is the config profile. In the config profile you should fill your Reddit API details.

For that please follow the steps below

```
git clone https://github.com/thecsw/prequelmemes_bot
cd prequelmemes_bot
mv example.config.py config.py
nano config.py
```

After filling out the details, save and exit. You're done with installation.

3 Deployment

Remove the word ***'example'*** from the title of all files with it.

Just run this

```
python main.py
```

That is everything. All the prequelmemes will be identified. Once and for all, the archives will never be incomplete.

4 Source code

The code is heavily commented and all the important modules are being separated into different files. Looks pretty, dunno.

5 Built With

1. praw is Python Reddit API Wrapper. This will be the main and only package to connect to Reddit's API and extract desired data.

2. python-opencv is used for image transformations and computer vision problems.

3. pytesseract is a python wrapper for Google's Tesseract-OCR.

4. tqdm is used for fancy progress bars.

5. Pillow is the Python Imaging Library by Fredrik Lundh and Contributors.

6 Authors

- **Sagindyk Urazayev** - *Initial work* - thecsw

- **farhank3389** - *Fixes* - farhank3389

7 License

This project is licensed under the The GNU General Public License (see the LICENSE.md file for details), it explains everything pretty well.

8   Todo list

This is going to be a short todo list. If you have the true loyalty and some free time, try to add/fix the following todo items or add your own with the issues tab or push.

8.1   TODO Edit the mesage body text.

The message body text isn't perfect. It can be more prequelmemed. So yeah, if somebody can do it, that would be awesome.

8.2   TODO Port the code to python3.6+

Just port the whole code. I am not talking about just using `python3 main.py` instead of `python main.py`. I am talking about using the new python3+ features. Like better string formatting, like the `print(f"{args}...")`.

8.3   TODO Get rid of the tqdm.

I don't know whym but tqdm has been throwing some errors at me. I should replace it with some **manual** loop counter. We are talking CS101 here. Heh

8.4   TODO Make some general fixes to the code.

Add some little new features, replace old ones with new ones. Make a faster version. Just a general quality assurance.

8.5   TODO Add a comment feature.

I don't have an idea of what to do. I made some prototypes, don't know. If the bot is going to answer to each comment, that is just a spam.