1. Generate a github token with access as required (Note: do not select admin, delete, write access for your token. Just keep it simple for reading purpose)
2. Place the token in file “fetch\_dataset\_features\_graduated\_projects.py”
3. Run “pip install PyGithub”
4. Remove the project urls from “graduated\_project\_updated\_urls.csv” and just keep those projects for which you are going to run the dataset collection.
5. Note the numbers of projects just so that others do not extract for the same project.
6. Run the script “fetch\_dataset\_features\_graduated\_projects.py” using the command “python fetch\_dataset\_features\_graduated\_projects.py”
7. Result will be populated in data\_set2.json file.