Gitter: https://gitter.im/scikit-learn/wimlds

Crash-Course in Contributing to Open Source Projects

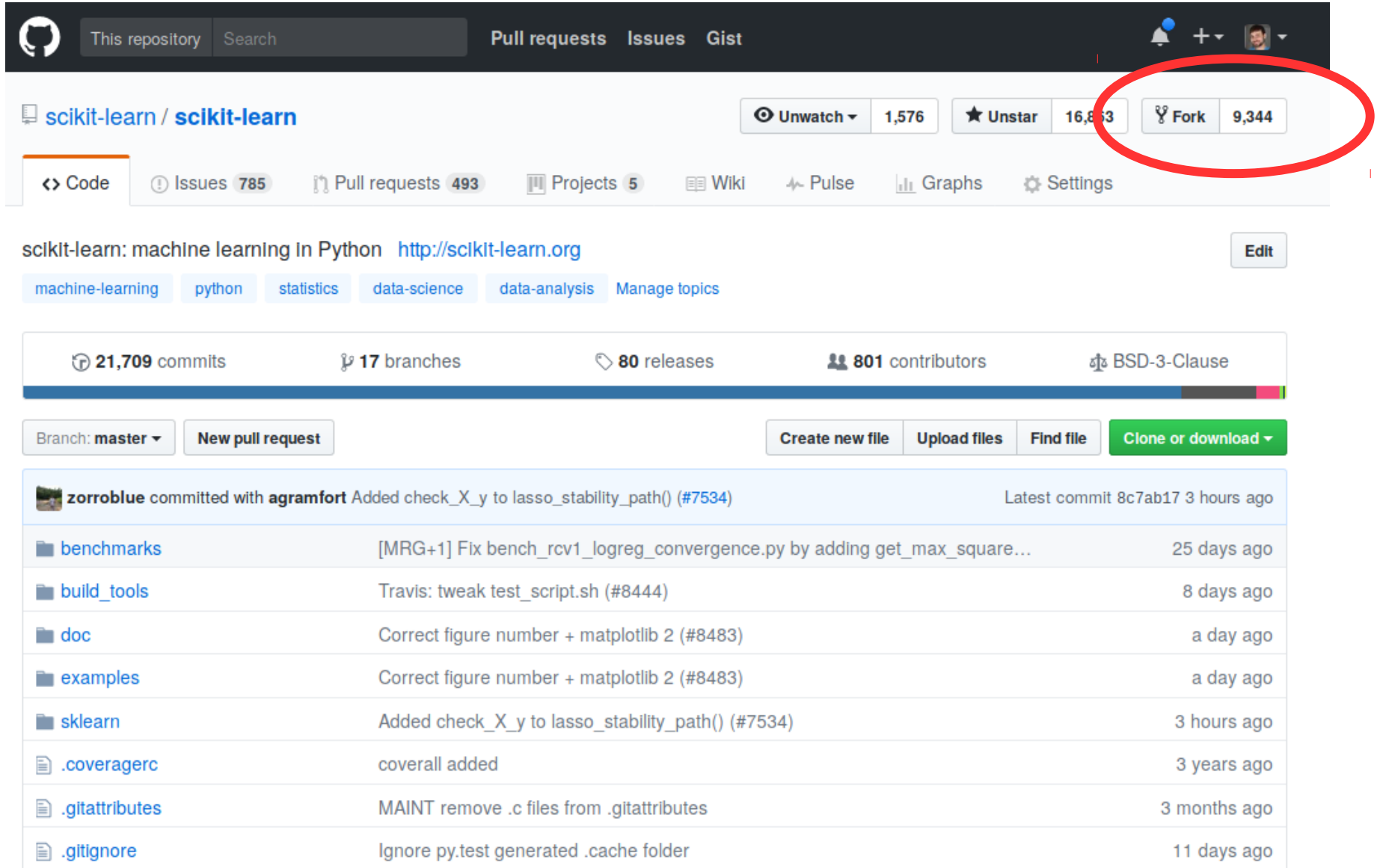# (with some specific instructions to sklearn)

Andreas Müller

Environment

- If you don't have a working python installation, install anaconda.

- If you don't want to install the scikit-learn development version in your main environment, create a separate environment (with conda or virtualenv):

```
conda create -n sklearndev numpy scipy matplotlib
     pytest sphinx cython ipykernel sphinxgallery
source activate sklearndev
```

Fork sklearn on github

https://github.com/scikit-learn/scikit-learn

Clone the Fork

If you don't have ssh keys set up, use https!
Add main scikit-learn repo as remote called "upstream":
`git remote add upstream git@github.com:scikit-learn/scikit-learn.git`

Build and run tests!

- Easiest way:
  ```
  cd scikit-learn
  pip install -e .
  ```

- Will overwrite existing installations!

Starting on Issues

- Go to https://github.com/scikit-learn/scikit-learn/issues?q=is%3Aissue+is%3Aopen+label%3ASprint

- Comment on issue saying "I'm working in this".

- Update local master branch from main sklearn repo (`git pull upstream master`).

- Create feature branch

  git checkout -b <branchname>

- Commit changes to branch, run tests

  `pytest sklearn`

  (or individual test files)

- Run flake8 on changed files

Creating a PR

- Push changes to your fork

  `git push origin <branchname>`

- Create PR using github UI:
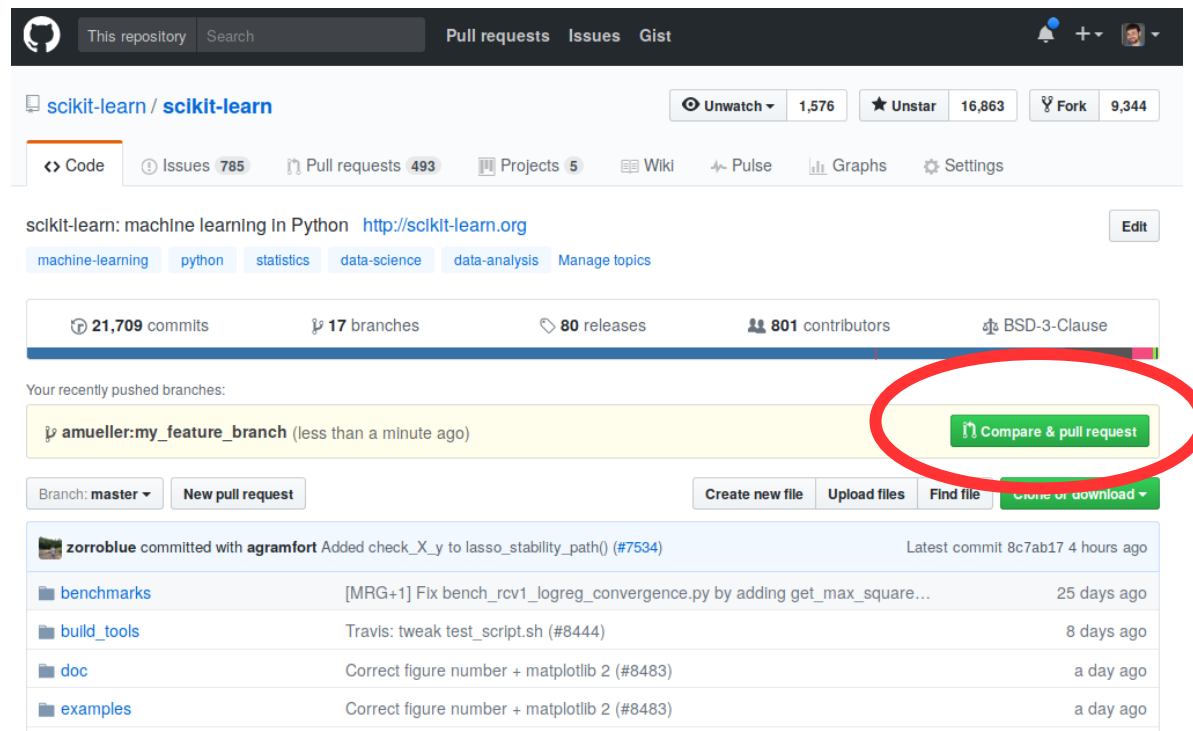
Describing PR

Open a pull request

Create a new pull request by comparing changes across two branches. If you need to, you can also compare across forks.

| base fork: **scikit-learn/scikit-learn** ▾ | base: **master** ▾ | ... | head fork: **amueller/scikit-learn** ▾ | compare: **my_feature_branch** ▾ | ✓ **Able to merge.** These br |

Please review the **guidelines for contributing** to this repository.
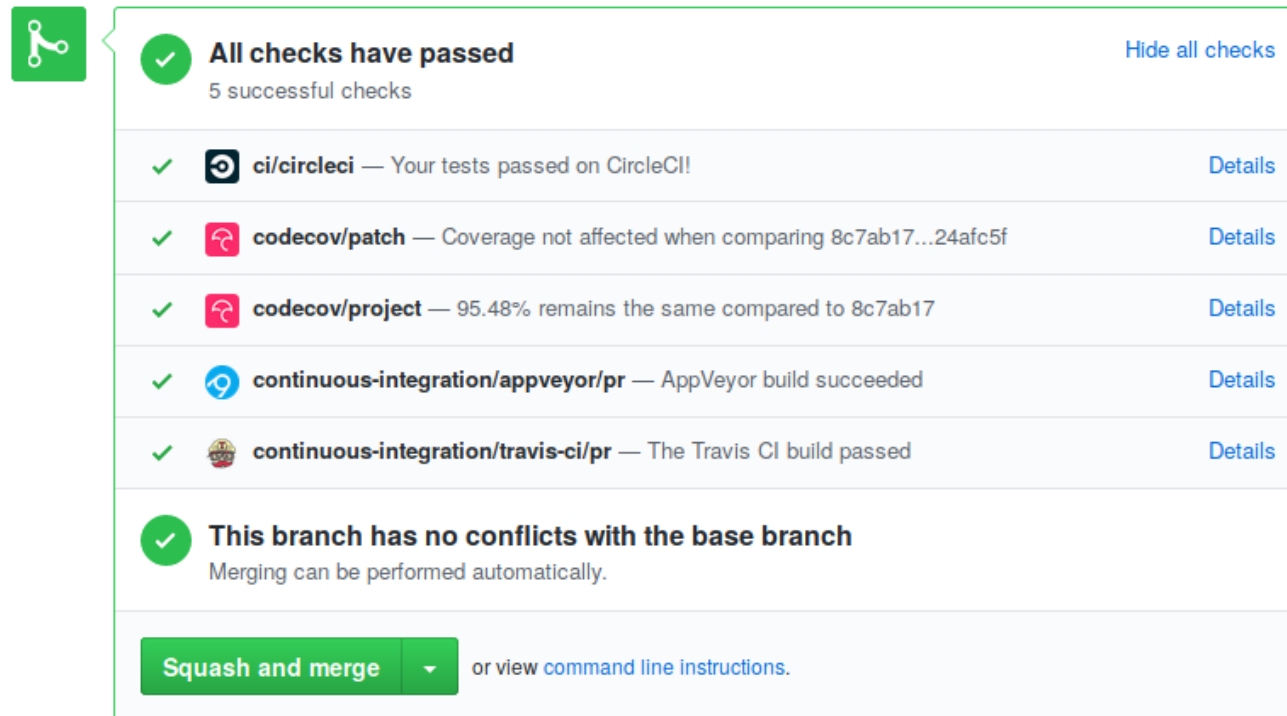
[MRG] minor pep8

[WIP] = work in progress
[MRG] = ready to merge

Write | Preview   AA ▾ B *i*   " <> ⊘   ☰ ☷ ☱   ↰ @ 🔖

```
<!--
Thanks for contributing a pull request! Please ensure you have taken a look at
the contribution guidelines: https://github.com/scikit-learn/scikit-learn/blob/master
/CONTRIBUTING.md#Contributing-Pull-Requests
-->
#### Reference Issue
<!-- Example: Fixes #1234 -->
```

Attach files by dragging & dropping, selecting them, or pasting from the clipboard.

☑ **Allow edits from maintainers.** Learn more

Create pull request

Reviewers ⚙
No reviews—request one

Assignees ⚙
No one—assign yourself

Labels ⚙
None yet

Milestone ⚙
No milestone

Regression tests

- Are mandatory! For everything (except documentation changes)!

- Make sure continuous integration passes:

What's next?

- Wait for reviews (be patient).
- Address review comments in the same branch.
- Pushing to your fork will update the PR
- Reviewers will "approve" PR or change title to [MRG + 1]
- You need two approvals for a merge.

Finding Issues

- Check "need contributor", "easy" and "sprint" issues.

- Something unclear in the docs? Fix it!

- Can't fix something that's unclear? Open an issue!

- Problem that you keep running into: Open an issue!

- Find stalled PRs (the author didn't address reviews for ~>1 month) and continue them!

- https://docs.google.com/document/d/1dgEPcGnbfs6DRs7qtZMkbF6mLP6cywfn46rMFGRmTQg/edit?usp=sharing

Reviewing

- You can review PRs and issues!
- Some bugs are not confirmed. See if you can confirm them and under what conditions?
- You can review documentation PRs for language and whether they are clear to you.
- You can review code changes on whether they address the issue (might be a bit tricky).
- Don't be afraid to ask clarifying questions!

Final words

- Pick something TRIVIALLY SIMPLE as the first contribution.

- You can do the cool stuff afterwards!

- There might be interesting issues that are not appropriately tagged.

Thank you for your help!
Enjoy!