MLOps

Continues Integration (CI)

Machine Learning Operations

Nicki Skafte Detlefsen,

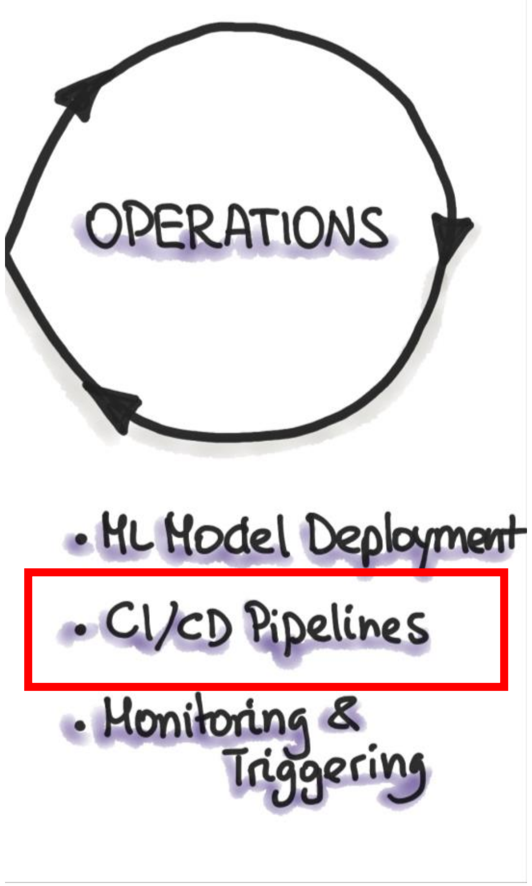
Postdoc

DTU Compute

Why you should care about todays exercises

What is continues integration?

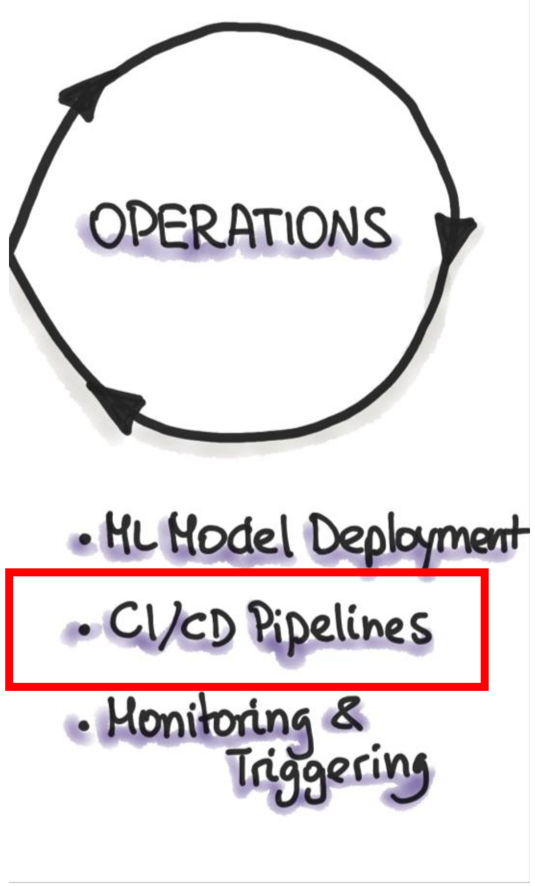
- Software practise
  - Frequently commit code to shared repository
  - By commit sooner than later, errors are captured early
  - Make merging easier
  - Automate build + test
- App independent



What is continues deployment?

- How to get your code to the user
  - 3 phases: Testing, staging and deploying
- App dependent

Note: Not covered in this course



What should you know about CI?

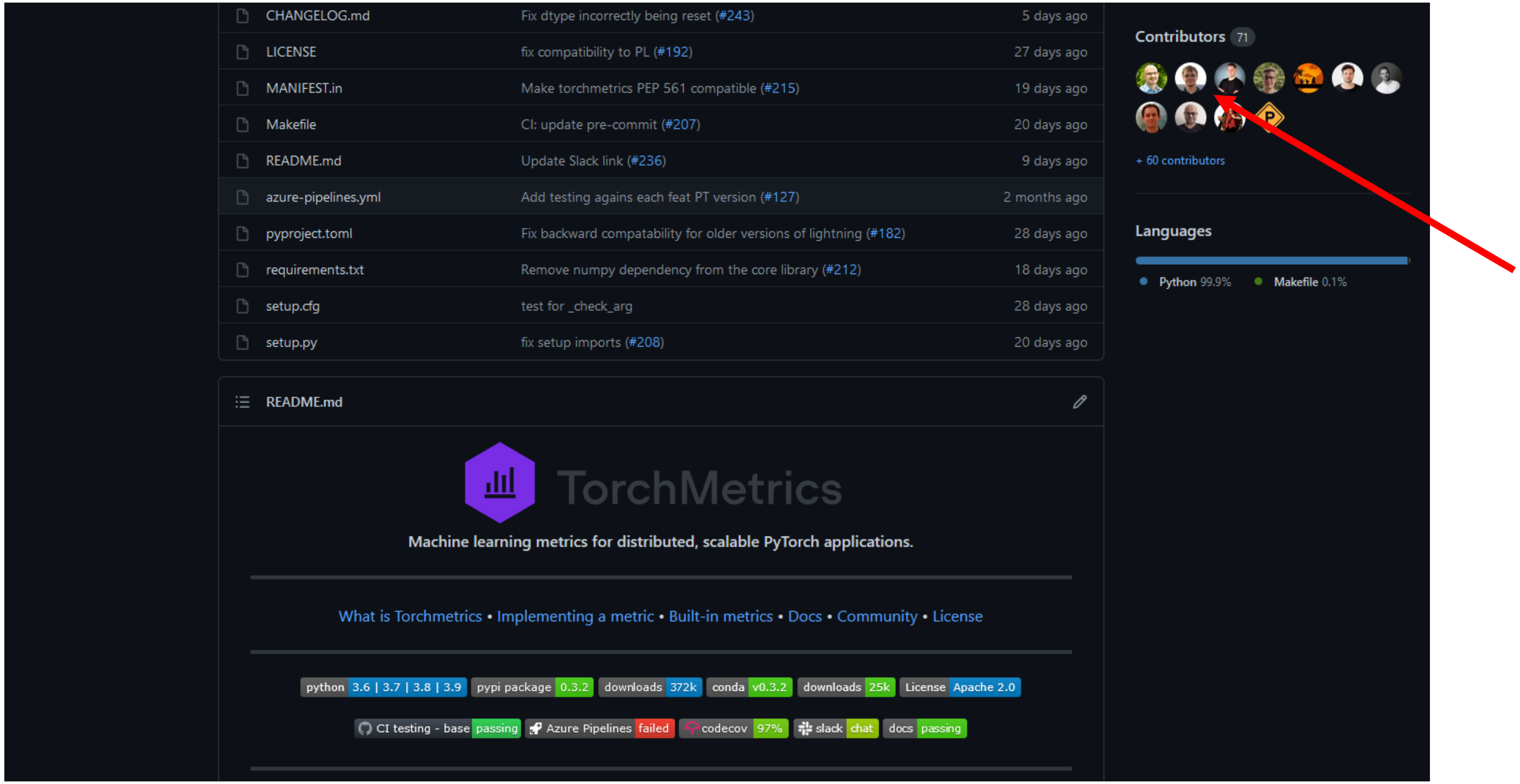
- CI is one of those topics that are best though as learning-by-doing

- If you understand how to use git, the rest is basically googling stuff

Let's look at a practical example

A small case study

All the metric in the world – now in pytorch



CHANGELOG.md	Fix dtype incorrectly being reset (#243)	5 days ago
LICENSE	fix compatibility to PL (#192)	27 days ago
MANIFEST.in	Make torchmetrics PEP 561 compatible (#215)	19 days ago
Makefile	CI: update pre-commit (#207)	20 days ago
README.md	Update Slack link (#236)	9 days ago
azure-pipelines.yml	Add testing agains each feat PT version (#127)	2 months ago
pyproject.toml	Fix backward compatability for older versions of lightning (#182)	28 days ago
requirements.txt	Remove numpy dependency from the core library (#212)	18 days ago
setup.cfg	test for _check_arg	28 days ago
setup.py	fix setup imports (#208)	20 days ago

README.md

TorchMetrics

Machine learning metrics for distributed, scalable PyTorch applications.

What is Torchmetrics • Implementing a metric • Built-in metrics • Docs • Community • License

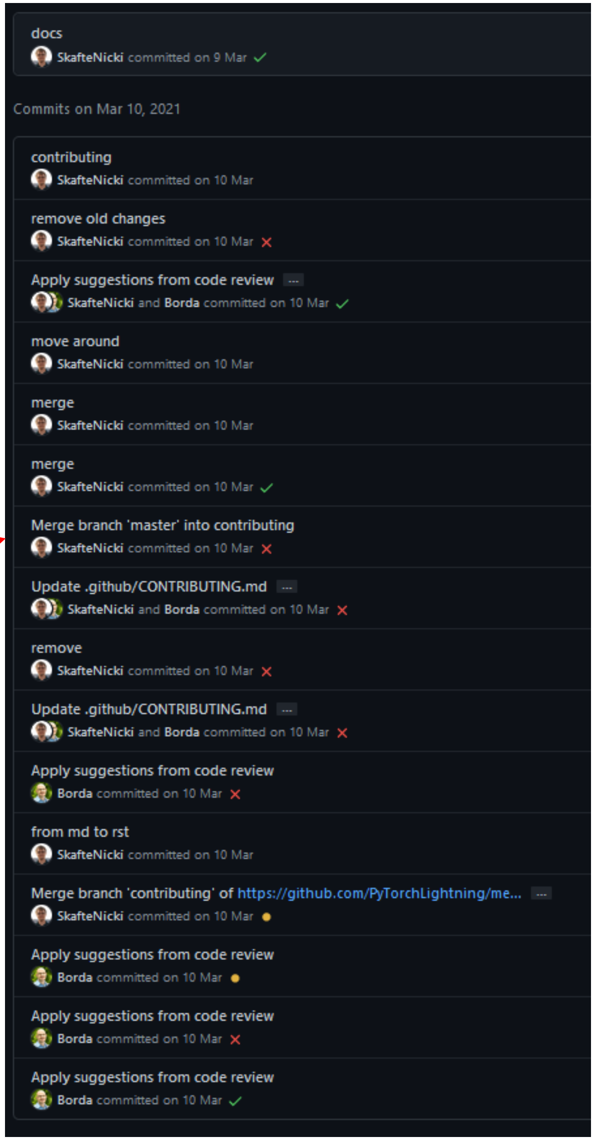
python 3.6 | 3.7 | 3.8 | 3.9 · pypi package 0.3.2 · downloads 372k · conda v0.3.2 · downloads 25k · License Apache 2.0

CI testing - base passing · Azure Pipelines failed · codecov 97% · slack chat · docs passing

Have I seen that guy before?

CI step 1: Commiting code

- Commit frequently
  - Catch errors sooner than later
  - Merging can be done automatically



docs
SkafteNicki committed on 9 Mar ✓

Commits on Mar 10, 2021

contributing
SkafteNicki committed on 10 Mar

remove old changes
SkafteNicki committed on 10 Mar ✗

Apply suggestions from code review
SkafteNicki and Borda committed on 10 Mar ✓

move around
SkafteNicki committed on 10 Mar

merge
SkafteNicki committed on 10 Mar

merge
SkafteNicki committed on 10 Mar ✓

Merge branch 'master' into contributing
SkafteNicki committed on 10 Mar ✗

Update .github/CONTRIBUTING.md
SkafteNicki and Borda committed on 10 Mar ✗

remove
SkafteNicki committed on 10 Mar ✗

Update .github/CONTRIBUTING.md
SkafteNicki and Borda committed on 10 Mar ✗

Apply suggestions from code review
Borda committed on 10 Mar ✗

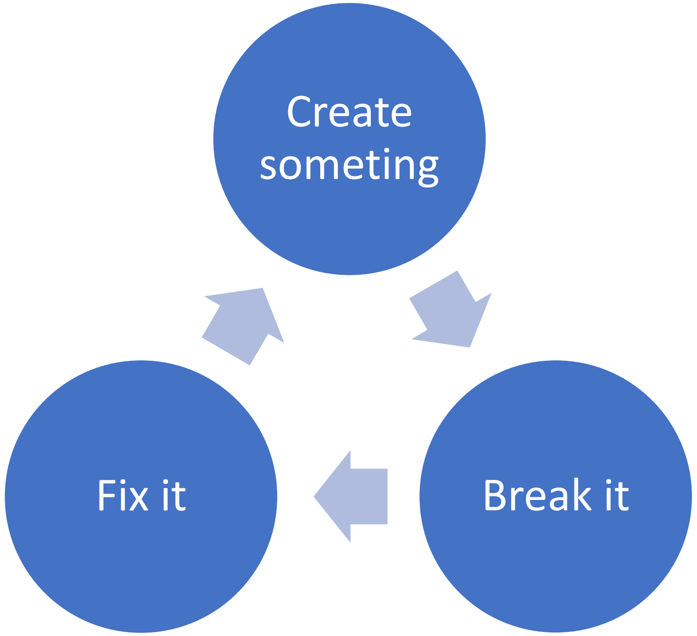
from md to rst
SkafteNicki committed on 10 Mar

Merge branch 'contributing' of https://github.com/PyTorchLightning/me...
SkafteNicki committed on 10 Mar ●

Apply suggestions from code review
Borda committed on 10 Mar ●

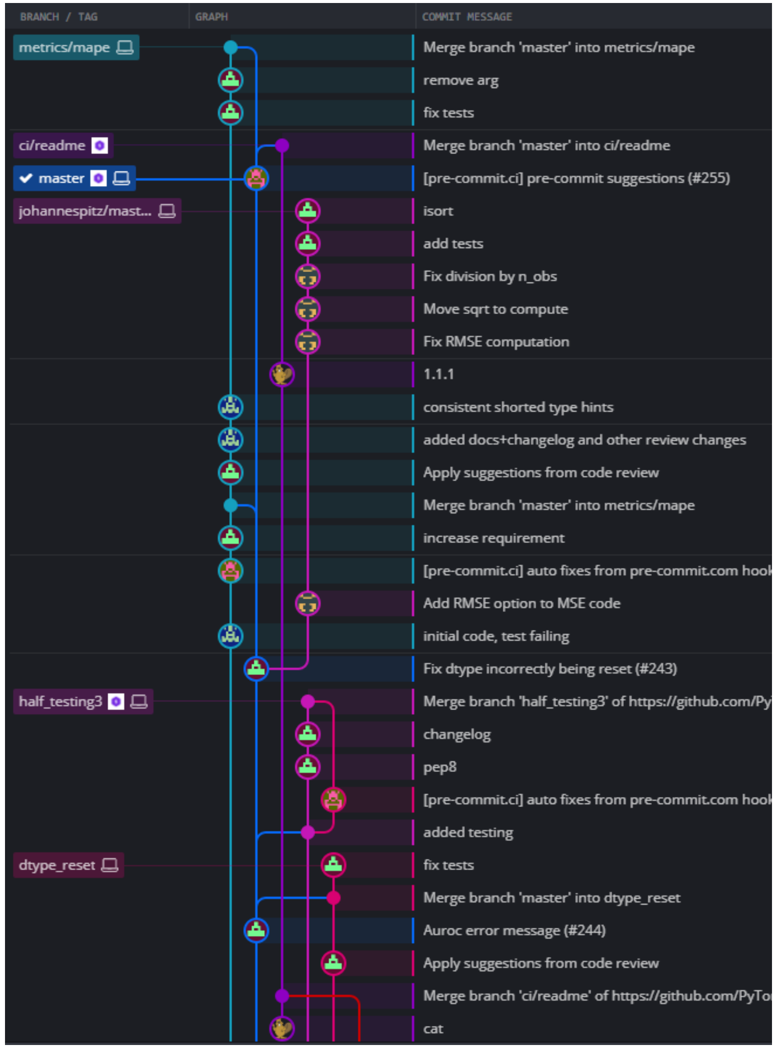
Apply suggestions from code review
Borda committed on 10 Mar ✗

Apply suggestions from code review
Borda committed on 10 Mar ✓



CI step 1: Commiting code

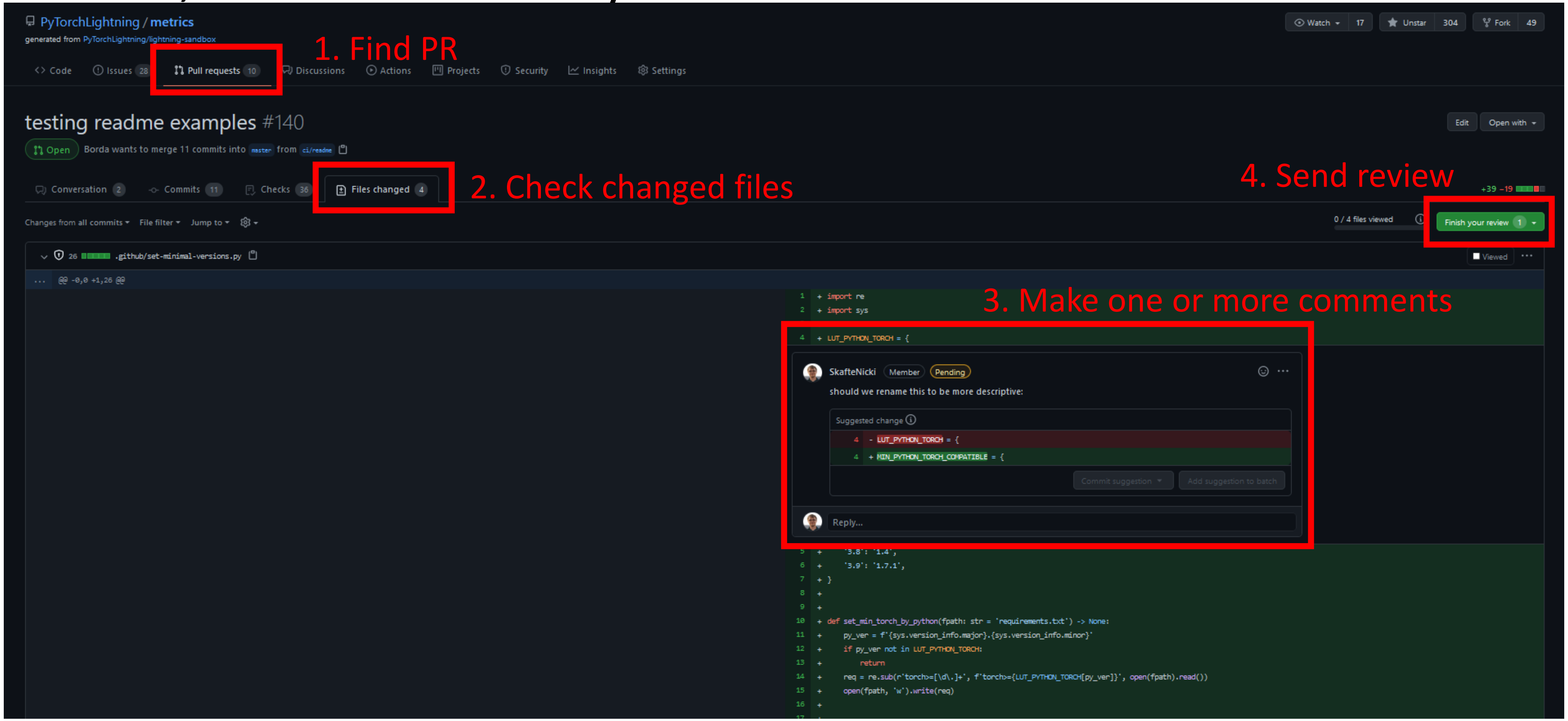
- Use branches
  - Enables parallel workflow
  - Experimental features/changes are keept away from ”stable” master



BRANCH / TAG	GRAPH	COMMIT MESSAGE
metrics/mape		Merge branch 'master' into metrics/mape
		remove arg
		fix tests
ci/readme		Merge branch 'master' into ci/readme
✓ master		[pre-commit.ci] pre-commit suggestions (#255)
johannespitz/mast...		isort
		add tests
		Fix division by n_obs
		Move sqrt to compute
		Fix RMSE computation
		1.1.1
		consistent shorted type hints
		added docs+changelog and other review changes
		Apply suggestions from code review
		Merge branch 'master' into metrics/mape
		increase requirement
		[pre-commit.ci] auto fixes from pre-commit.com hook
		Add RMSE option to MSE code
		initial code, test failing
		Fix dtype incorrectly being reset (#243)
half_testing3		Merge branch 'half_testing3' of https://github.com/Py
		changelog
		pep8
		[pre-commit.ci] auto fixes from pre-commit.com hook
		added testing
dtype_reset		fix tests
		Merge branch 'master' into dtype_reset
		Auroc error message (#244)
		Apply suggestions from code review
		Merge branch 'ci/readme' of https://github.com/PyTo
		cat

CI step 1: Commiting code

- Use PRs, other can review your code



1. Find PR

2. Check changed files

3. Make one or more comments

4. Send review

testing readme examples #140

Open — Borda wants to merge 11 commits into `master` from `ci/readme`

26 .github/set-minimal-versions.py

```
@@ -0,0 +1,26 @@
+ import re
+ import sys
+ LUT_PYTHON_TORCH = {
```

SkafteNicki: should we rename this to be more descriptive:

Suggested change

```
- LUT_PYTHON_TORCH = {
+ MIN_PYTHON_TORCH_COMPATIBLE = {
```

```
+     '3.8': '1.4',
+     '3.9': '1.7.1',
+ }
+
+
+ def set_min_torch_by_python(fpath: str = 'requirements.txt') -> None:
+     py_ver = f'{sys.version_info.major}.{sys.version_info.minor}'
+     if py_ver not in LUT_PYTHON_TORCH:
+         return
+     req = re.sub(r'torch>=[\d\.]+', f'torch>={LUT_PYTHON_TORCH[py_ver]}', open(fpath).read())
+     open(fpath, 'w').write(req)
+
```

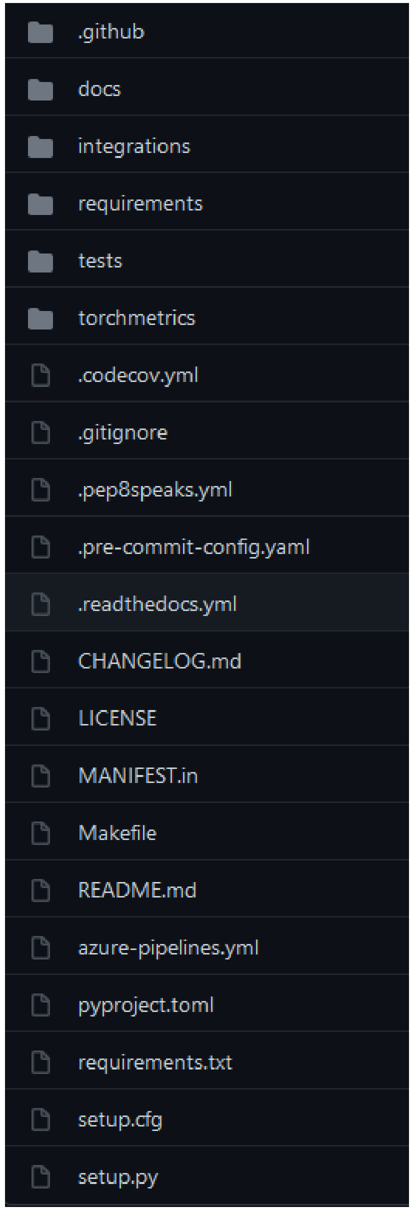
CI step 2: Automating stuff

- What can be automated: EVERYTHING
  - Functional tests
  - Documentation creation
  - Linters (which check style formatting)
  - Security checks
  - Code coverage
  - Custom checks…

A small casestudy

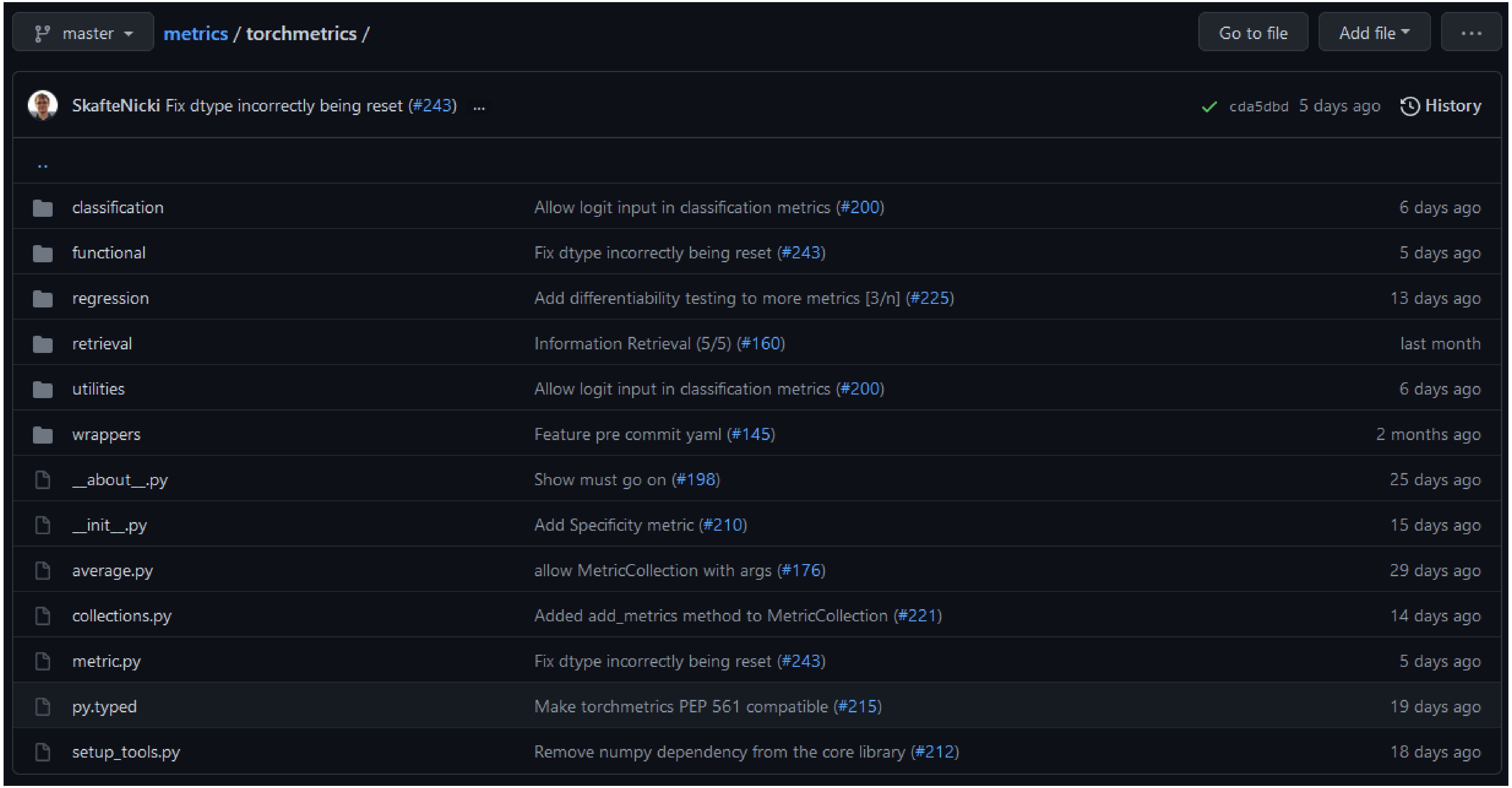
Ranked by importance (biased)

1. Source code
2. Tests
3. Setup
4. CI workflows
5. Documentation
6. ★



- .github
- docs
- integrations
- requirements
- tests
- torchmetrics
- .codecov.yml
- ★ .gitignore
- .pep8speaks.yml
- .pre-commit-config.yaml
- .readthedocs.yml
- CHANGELOG.md
- ★ LICENSE
- MANIFEST.in
- Makefile
- ★ README.md
- azure-pipelines.yml
- pyproject.toml
- ★ requirements.txt
- setup.cfg
- setup.py

Source code



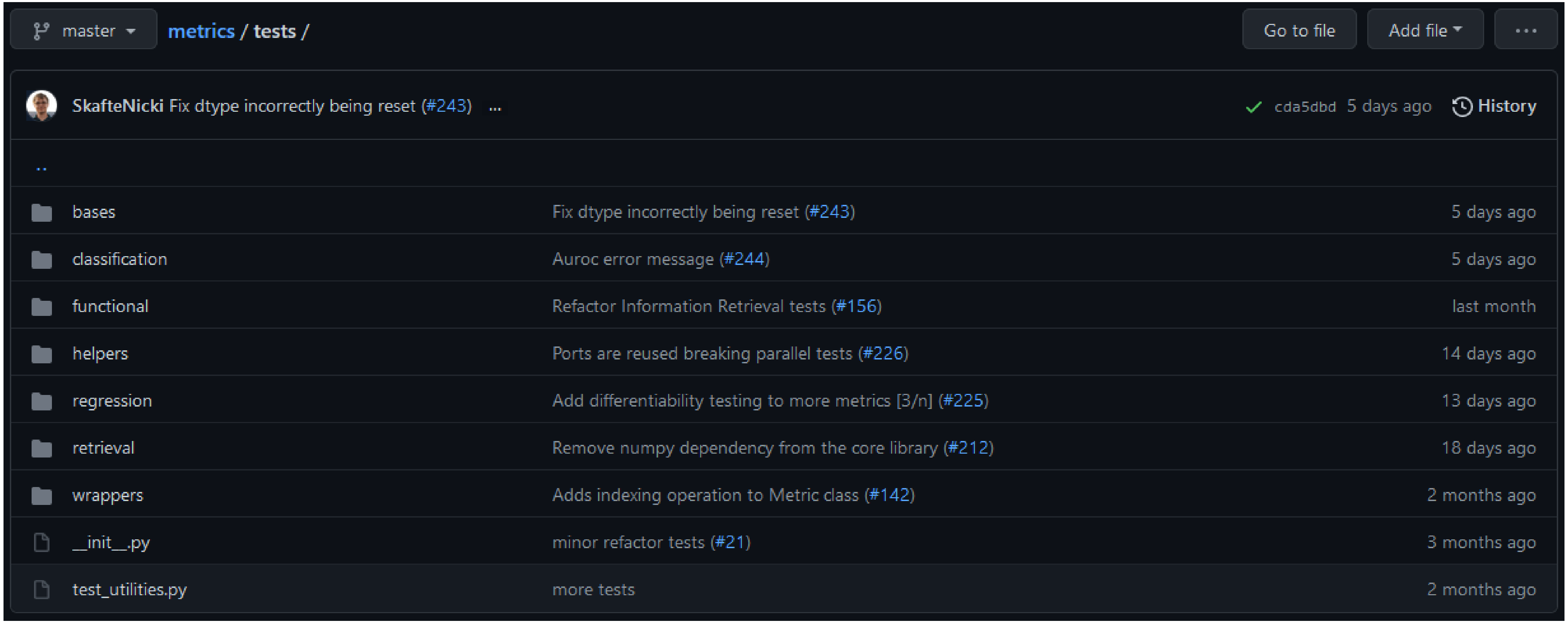
SkafteNicki Fix dtype incorrectly being reset (#243) … ✓ cda5dbd 5 days ago History

..

classification	Allow logit input in classification metrics (#200)	6 days ago
functional	Fix dtype incorrectly being reset (#243)	5 days ago
regression	Add differentiability testing to more metrics [3/n] (#225)	13 days ago
retrieval	Information Retrieval (5/5) (#160)	last month
utilities	Allow logit input in classification metrics (#200)	6 days ago
wrappers	Feature pre commit yaml (#145)	2 months ago
__about__.py	Show must go on (#198)	25 days ago
__init__.py	Add Specificity metric (#210)	15 days ago
average.py	allow MetricCollection with args (#176)	29 days ago
collections.py	Added add_metrics method to MetricCollection (#221)	14 days ago
metric.py	Fix dtype incorrectly being reset (#243)	5 days ago
py.typed	Make torchmetrics PEP 561 compatible (#215)	19 days ago
setup_tools.py	Remove numpy dependency from the core library (#212)	18 days ago

Test code

In total we have 6365 tests



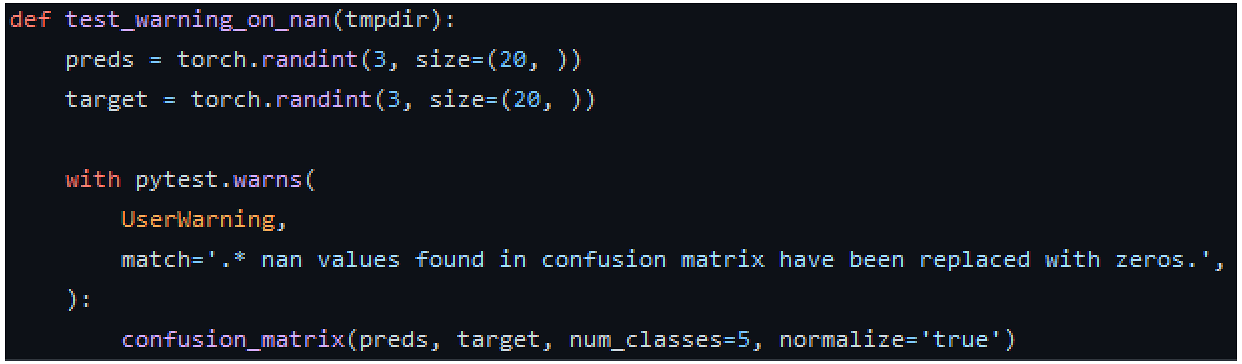
SkafteNicki Fix dtype incorrectly being reset (#243) ... ✓ cda5dbd 5 days ago

..

bases	Fix dtype incorrectly being reset (#243)	5 days ago
classification	Auroc error message (#244)	5 days ago
functional	Refactor Information Retrieval tests (#156)	last month
helpers	Ports are reused breaking parallel tests (#226)	14 days ago
regression	Add differentiability testing to more metrics [3/n] (#225)	13 days ago
retrieval	Remove numpy dependency from the core library (#212)	18 days ago
wrappers	Adds indexing operation to Metric class (#142)	2 months ago
__init__.py	minor refactor tests (#21)	3 months ago
test_utilities.py	more tests	2 months ago

Test example 1

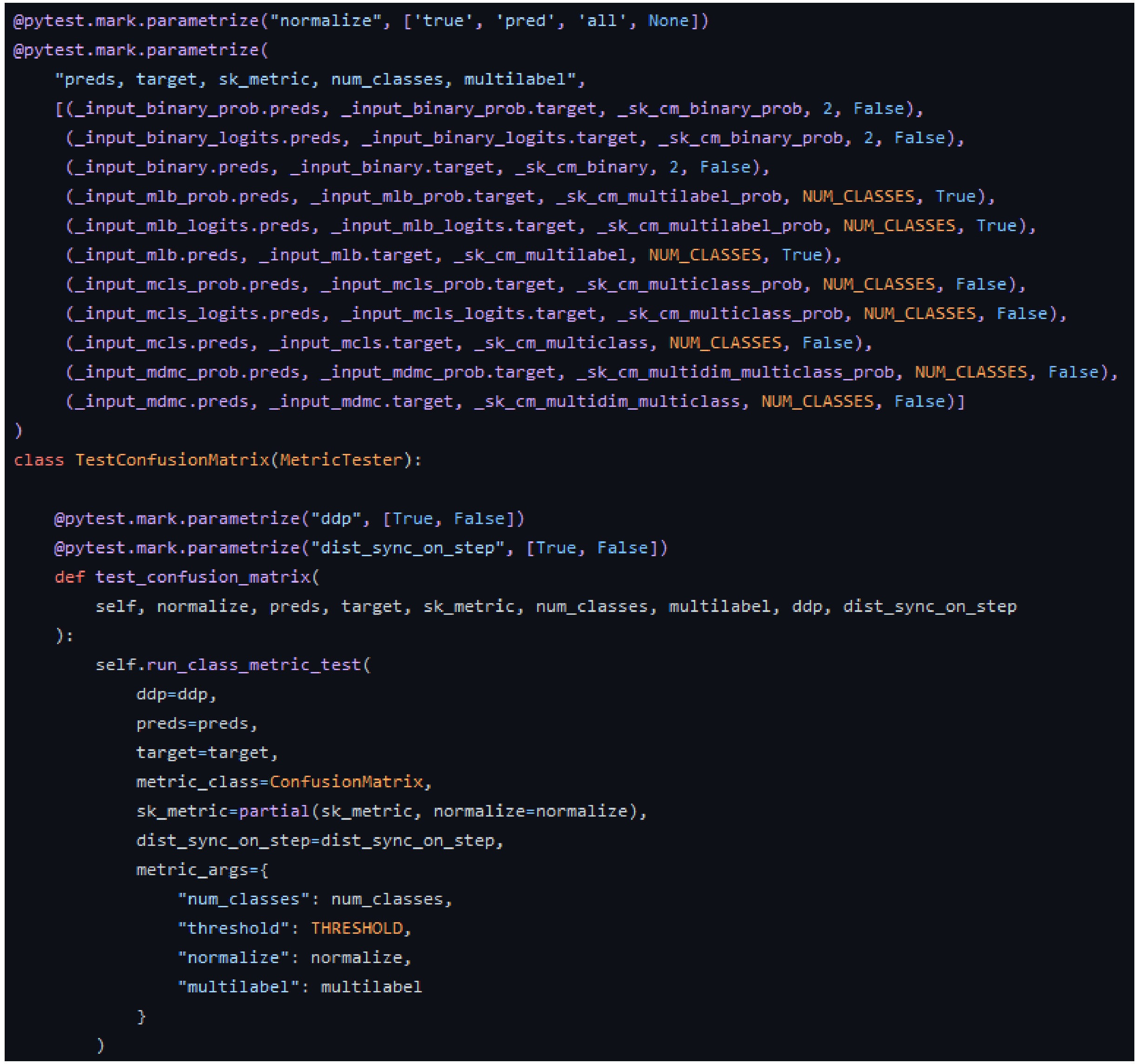
Can be simple



```
def test_warning_on_nan(tmpdir):
    preds = torch.randint(3, size=(20, ))
    target = torch.randint(3, size=(20, ))

    with pytest.warns(
        UserWarning,
        match='.* nan values found in confusion matrix have been replaced with zeros.',
    ):
        confusion_matrix(preds, target, num_classes=5, normalize='true')
```

Test example 2

Can be very complex



```
@pytest.mark.parametrize("normalize", ['true', 'pred', 'all', None])
@pytest.mark.parametrize(
    "preds, target, sk_metric, num_classes, multilabel",
    [(_input_binary_prob.preds, _input_binary_prob.target, _sk_cm_binary_prob, 2, False),
     (_input_binary_logits.preds, _input_binary_logits.target, _sk_cm_binary_prob, 2, False),
     (_input_binary.preds, _input_binary.target, _sk_cm_binary, 2, False),
     (_input_mlb_prob.preds, _input_mlb_prob.target, _sk_cm_multilabel_prob, NUM_CLASSES, True),
     (_input_mlb_logits.preds, _input_mlb_logits.target, _sk_cm_multilabel_prob, NUM_CLASSES, True),
     (_input_mlb.preds, _input_mlb.target, _sk_cm_multilabel, NUM_CLASSES, True),
     (_input_mcls_prob.preds, _input_mcls_prob.target, _sk_cm_multiclass_prob, NUM_CLASSES, False),
     (_input_mcls_logits.preds, _input_mcls_logits.target, _sk_cm_multiclass_prob, NUM_CLASSES, False),
     (_input_mcls.preds, _input_mcls.target, _sk_cm_multiclass, NUM_CLASSES, False),
     (_input_mdmc_prob.preds, _input_mdmc_prob.target, _sk_cm_multidim_multiclass_prob, NUM_CLASSES, False),
     (_input_mdmc.preds, _input_mdmc.target, _sk_cm_multidim_multiclass, NUM_CLASSES, False)]
)
class TestConfusionMatrix(MetricTester):

    @pytest.mark.parametrize("ddp", [True, False])
    @pytest.mark.parametrize("dist_sync_on_step", [True, False])
    def test_confusion_matrix(
        self, normalize, preds, target, sk_metric, num_classes, multilabel, ddp, dist_sync_on_step
    ):
        self.run_class_metric_test(
            ddp=ddp,
            preds=preds,
            target=target,
            metric_class=ConfusionMatrix,
            sk_metric=partial(sk_metric, normalize=normalize),
            dist_sync_on_step=dist_sync_on_step,
            metric_args={
                "num_classes": num_classes,
                "threshold": THRESHOLD,
                "normalize": normalize,
                "multilabel": multilabel
            }
        )
```

Setup files

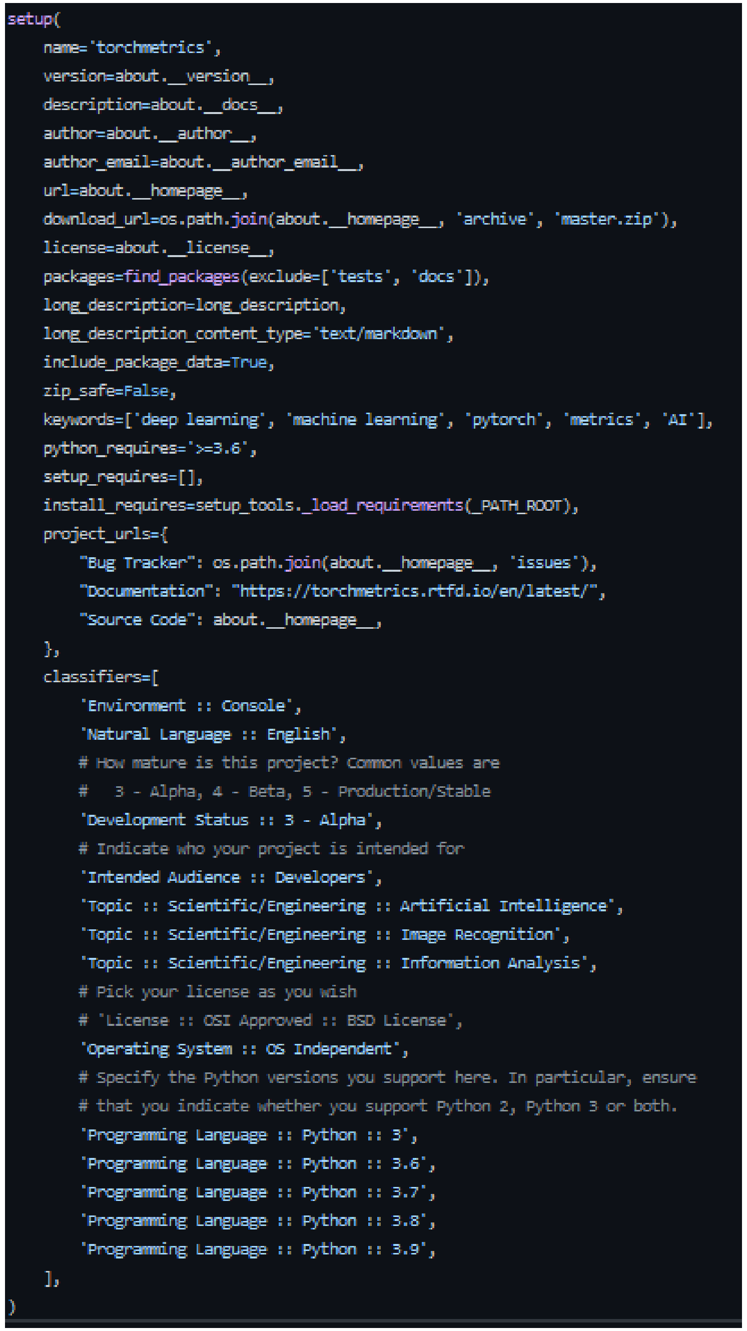
- Contains all information regarding the project

- Allow people to do:

python setup.py install

or if uploaded to pip

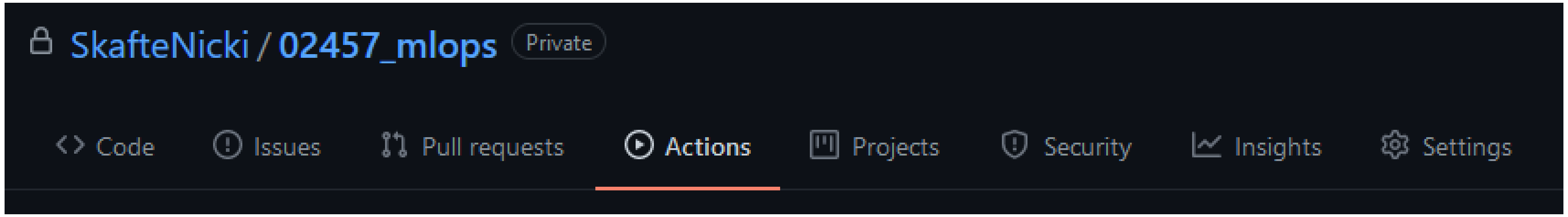
pip install my_package



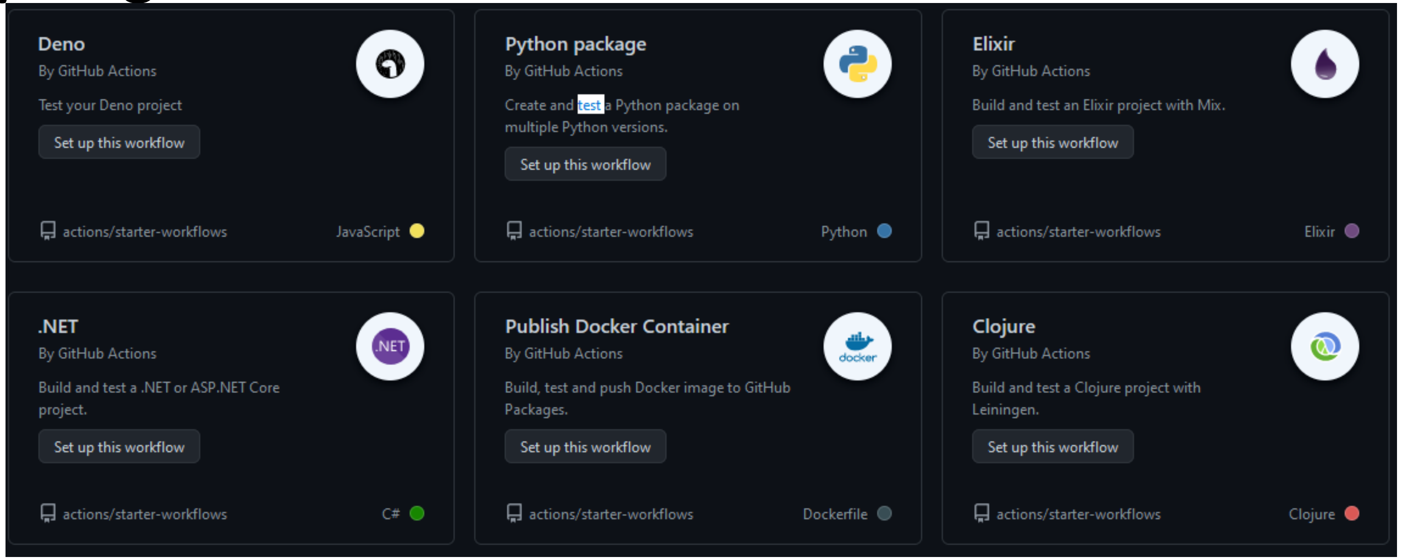
```
setup(
    name='torchmetrics',
    version=about.__version__,
    description=about.__docs__,
    author=about.__author__,
    author_email=about.__author_email__,
    url=about.__homepage__,
    download_url=os.path.join(about.__homepage__, 'archive', 'master.zip'),
    license=about.__license__,
    packages=find_packages(exclude=['tests', 'docs']),
    long_description=long_description,
    long_description_content_type='text/markdown',
    include_package_data=True,
    zip_safe=False,
    keywords=['deep learning', 'machine learning', 'pytorch', 'metrics', 'AI'],
    python_requires='>=3.6',
    setup_requires=[],
    install_requires=setup_tools._load_requirements(_PATH_ROOT),
    project_urls={
        "Bug Tracker": os.path.join(about.__homepage__, 'issues'),
        "Documentation": "https://torchmetrics.rtfd.io/en/latest/",
        "Source Code": about.__homepage__,
    },
    classifiers=[
        'Environment :: Console',
        'Natural Language :: English',
        # How mature is this project? Common values are
        #   3 - Alpha, 4 - Beta, 5 - Production/Stable
        'Development Status :: 3 - Alpha',
        # Indicate who your project is intended for
        'Intended Audience :: Developers',
        'Topic :: Scientific/Engineering :: Artificial Intelligence',
        'Topic :: Scientific/Engineering :: Image Recognition',
        'Topic :: Scientific/Engineering :: Information Analysis',
        # Pick your license as you wish
        # 'License :: OSI Approved :: BSD License',
        'Operating System :: OS Independent',
        # Specify the Python versions you support here. In particular, ensure
        # that you indicate whether you support Python 2, Python 3 or both.
        'Programming Language :: Python :: 3',
        'Programming Language :: Python :: 3.6',
        'Programming Language :: Python :: 3.7',
        'Programming Language :: Python :: 3.8',
        'Programming Language :: Python :: 3.9',
    ],
)
```

Github actions

- Build-in CI for github
- Free 2,000 automation minutes/month (public repository)



- Many ready to go workflows



Deno
By GitHub Actions
Test your Deno project
Set up this workflow
actions/starter-workflows JavaScript

Python package
By GitHub Actions
Create and test a Python package on multiple Python versions.
Set up this workflow
actions/starter-workflows Python

Elixir
By GitHub Actions
Build and test an Elixir project with Mix.
Set up this workflow
actions/starter-workflows Elixir

.NET
By GitHub Actions
Build and test a .NET or ASP.NET Core project.
Set up this workflow
actions/starter-workflows C#

Publish Docker Container
By GitHub Actions
Build, test and push Docker image to GitHub Packages.
Set up this workflow
actions/starter-workflows Dockerfile

Clojure
By GitHub Actions
Build and test a Clojure project with Leiningen.
Set up this workflow
actions/starter-workflows Clojure

The anatomy of a workflow file

When tests should be triggered

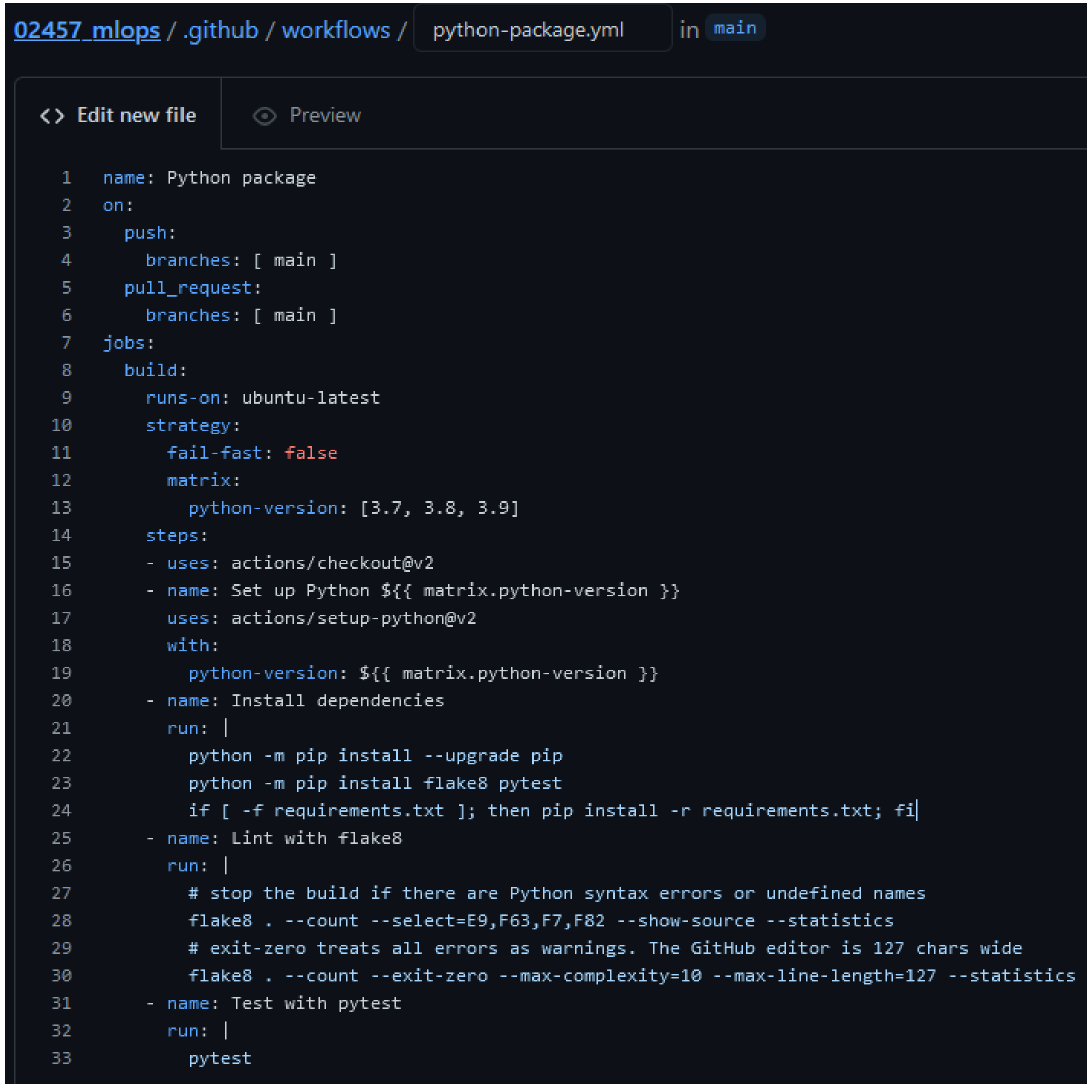
Define Operating system + python version

Setup python

Install dependencies

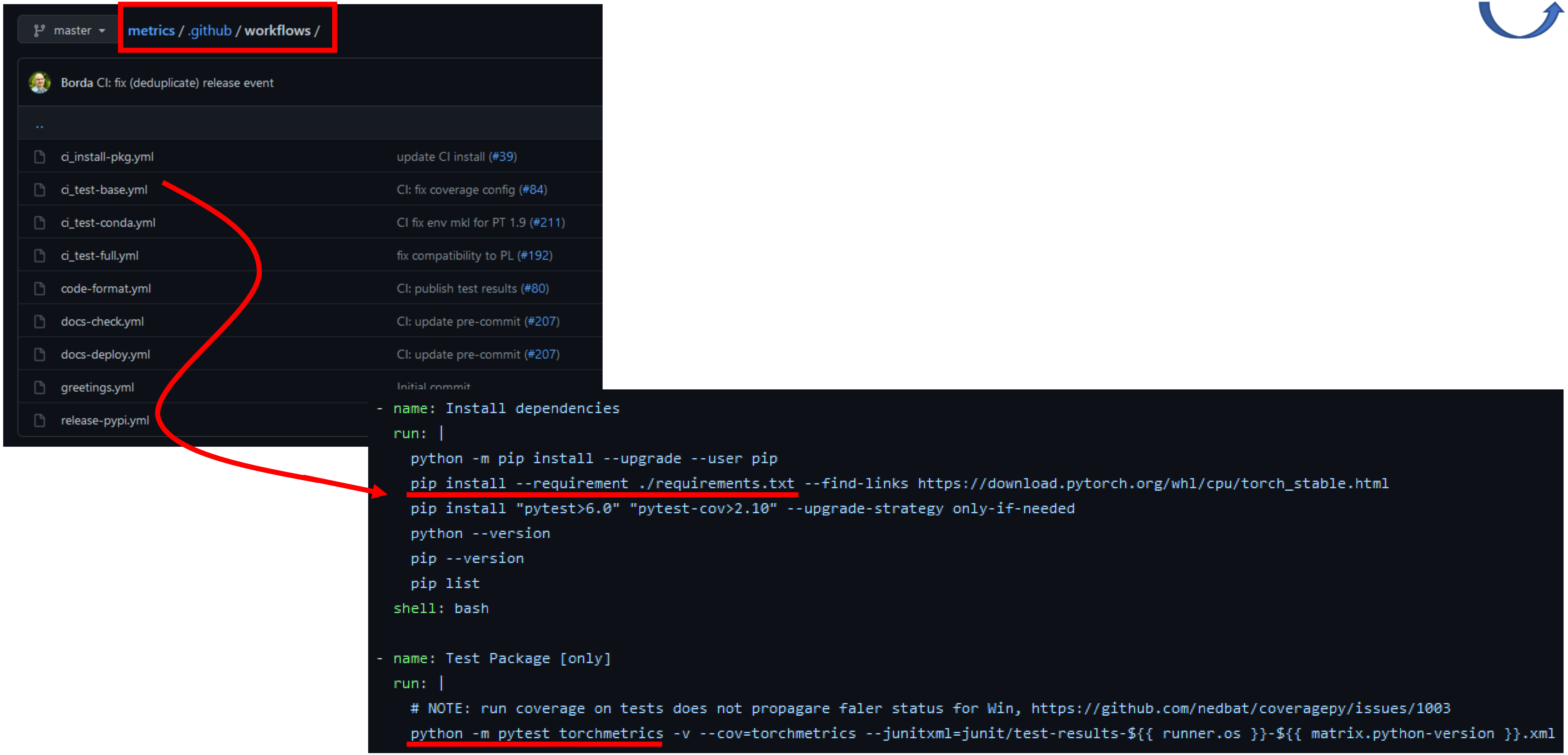
Check linting (stop if errors)

Run tests


02457_mlops / .github / workflows / python-package.yml in main
Edit new file
Preview

```
name: Python package
on:
  push:
    branches: [ main ]
  pull_request:
    branches: [ main ]
jobs:
  build:
    runs-on: ubuntu-latest
    strategy:
      fail-fast: false
      matrix:
        python-version: [3.7, 3.8, 3.9]
    steps:
    - uses: actions/checkout@v2
    - name: Set up Python ${{ matrix.python-version }}
      uses: actions/setup-python@v2
      with:
        python-version: ${{ matrix.python-version }}
    - name: Install dependencies
      run: |
        python -m pip install --upgrade pip
        python -m pip install flake8 pytest
        if [ -f requirements.txt ]; then pip install -r requirements.txt; fi
    - name: Lint with flake8
      run: |
        # stop the build if there are Python syntax errors or undefined names
        flake8 . --count --select=E9,F63,F7,F82 --show-source --statistics
        # exit-zero treats all errors as warnings. The GitHub editor is 127 chars wide
        flake8 . --count --exit-zero --max-complexity=10 --max-line-length=127 --statistics
    - name: Test with pytest
      run: |
        pytest
```

A small case study



master

metrics / .github / workflows /

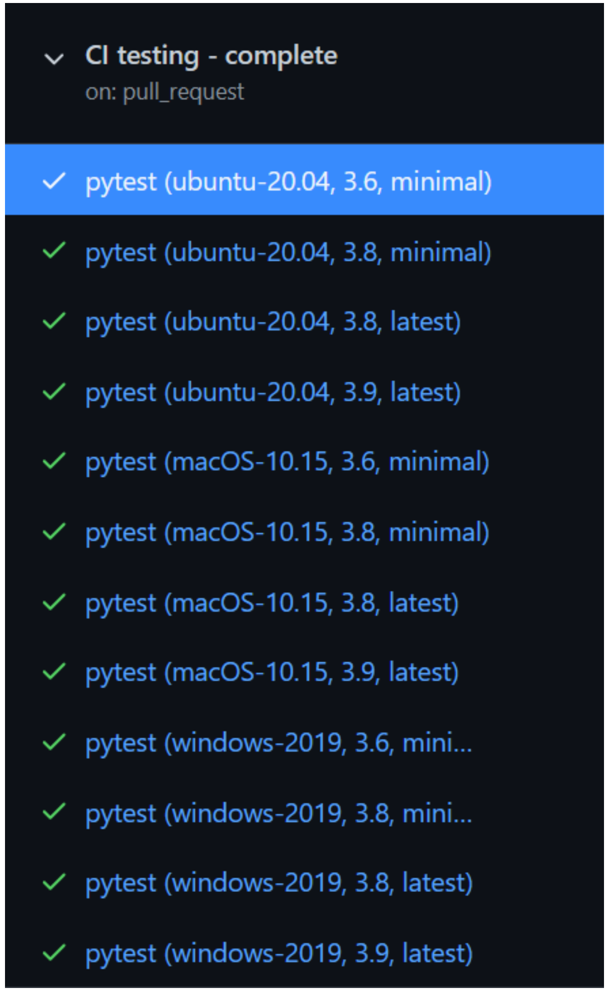
Borda CI: fix (deduplicate) release event

..

ci_install-pkg.yml	update CI install (#39)
ci_test-base.yml	CI: fix coverage config (#84)
ci_test-conda.yml	CI fix env mkl for PT 1.9 (#211)
ci_test-full.yml	fix compatibility to PL (#192)
code-format.yml	CI: publish test results (#80)
docs-check.yml	CI: update pre-commit (#207)
docs-deploy.yml	CI: update pre-commit (#207)
greetings.yml	Initial commit
release-pypi.yml	

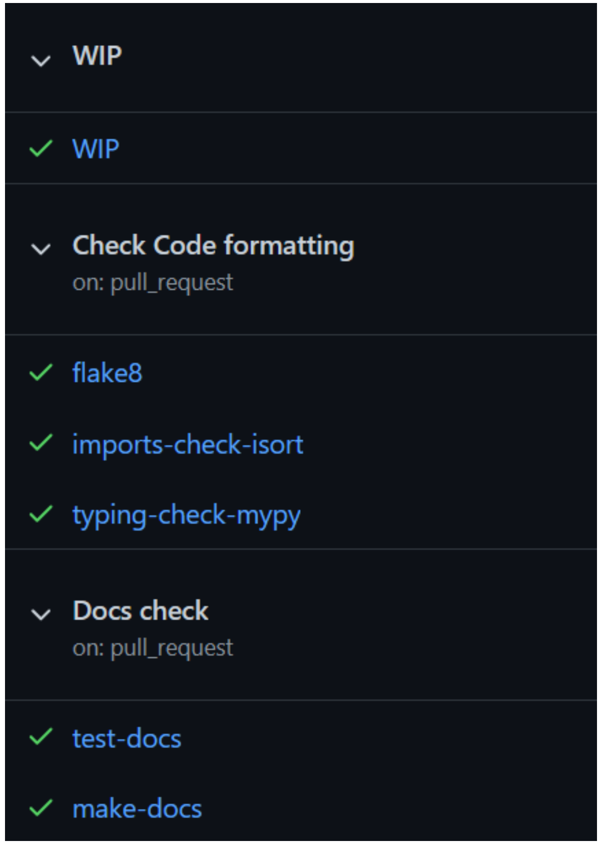
```
- name: Install dependencies
  run: |
    python -m pip install --upgrade --user pip
    pip install --requirement ./requirements.txt --find-links https://download.pytorch.org/whl/cpu/torch_stable.html
    pip install "pytest>6.0" "pytest-cov>2.10" --upgrade-strategy only-if-needed
    python --version
    pip --version
    pip list
  shell: bash

- name: Test Package [only]
  run: |
    # NOTE: run coverage on tests does not propagare faler status for Win, https://github.com/nedbat/coveragepy/issues/1003
    python -m pytest torchmetrics -v --cov=torchmetrics --junitxml=junit/test-results-${{ runner.os }}-${{ matrix.python-version }}.xml
```

All of our workflows – 36 in total



CI testing - complete
on: pull_request

- ✓ pytest (ubuntu-20.04, 3.6, minimal)
- ✓ pytest (ubuntu-20.04, 3.8, minimal)
- ✓ pytest (ubuntu-20.04, 3.8, latest)
- ✓ pytest (ubuntu-20.04, 3.9, latest)
- ✓ pytest (macOS-10.15, 3.6, minimal)
- ✓ pytest (macOS-10.15, 3.8, minimal)
- ✓ pytest (macOS-10.15, 3.8, latest)
- ✓ pytest (macOS-10.15, 3.9, latest)
- ✓ pytest (windows-2019, 3.6, mini...
- ✓ pytest (windows-2019, 3.8, mini...
- ✓ pytest (windows-2019, 3.8, latest)
- ✓ pytest (windows-2019, 3.9, latest)



WIP

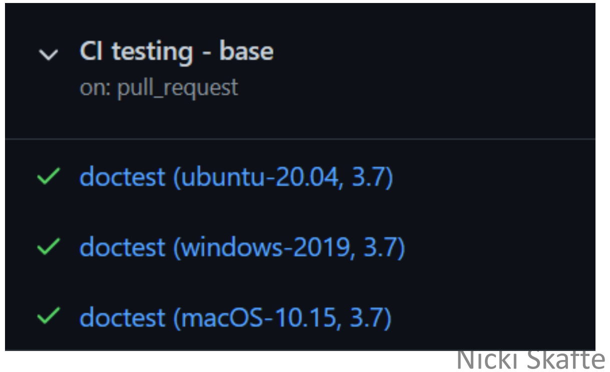
- ✓ WIP

Check Code formatting
on: pull_request

- ✓ flake8
- ✓ imports-check-isort
- ✓ typing-check-mypy

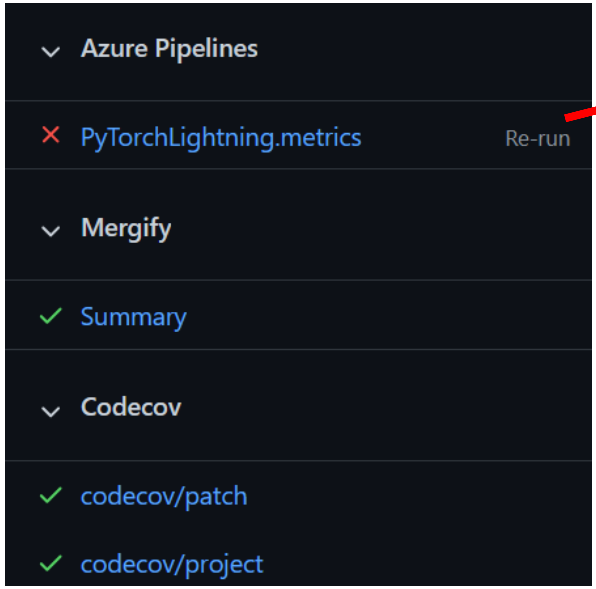
Docs check
on: pull_request

- ✓ test-docs
- ✓ make-docs



CI testing - base
on: pull_request

- ✓ doctest (ubuntu-20.04, 3.7)
- ✓ doctest (windows-2019, 3.7)
- ✓ doctest (macOS-10.15, 3.7)



Azure Pipelines

- ✗ PyTorchLightning.metrics Re-run

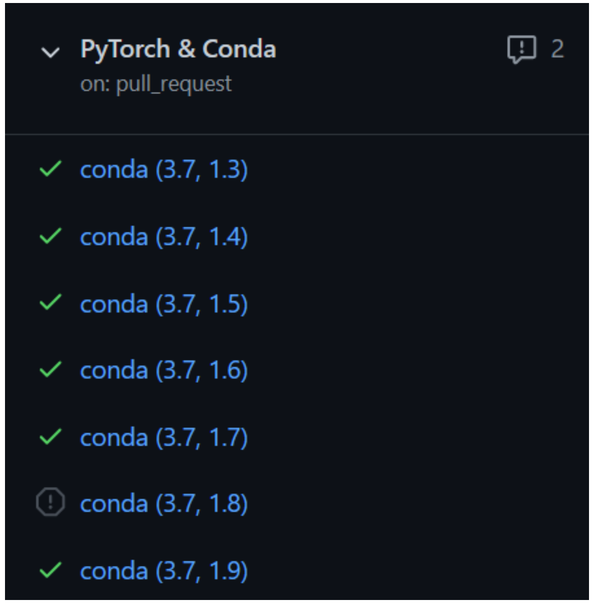
Mergify

- ✓ Summary

Codecov

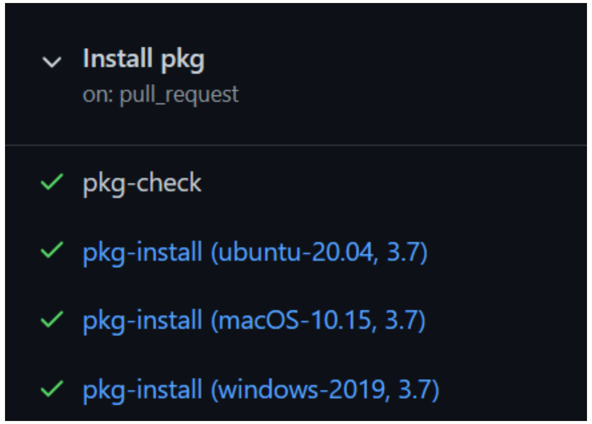
- ✓ codecov/patch
- ✓ codecov/project

Not github actions



PyTorch & Conda 2
on: pull_request

- ✓ conda (3.7, 1.3)
- ✓ conda (3.7, 1.4)
- ✓ conda (3.7, 1.5)
- ✓ conda (3.7, 1.6)
- ✓ conda (3.7, 1.7)
- ! conda (3.7, 1.8)
- ✓ conda (3.7, 1.9)



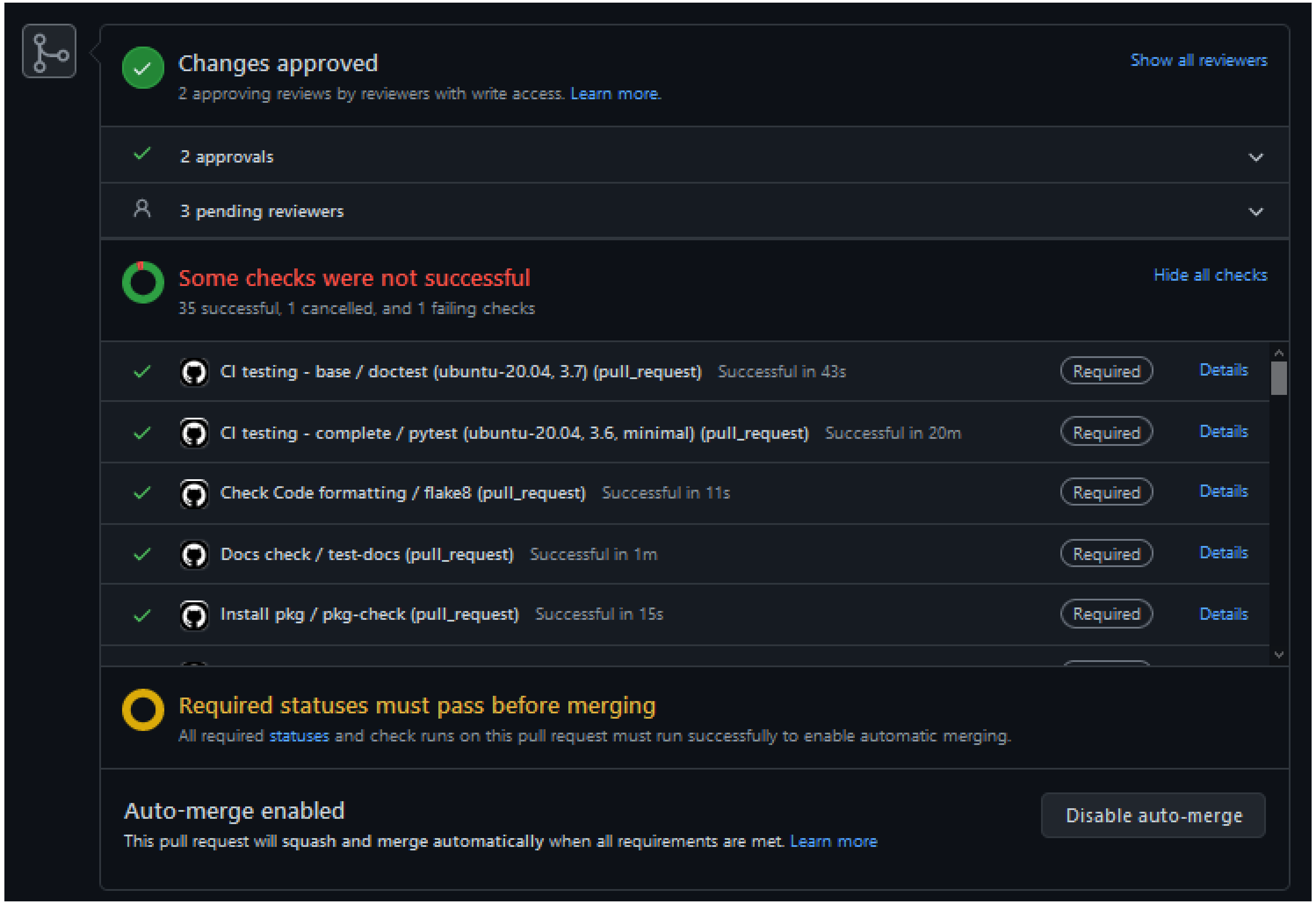
Install pkg
on: pull_request

- ✓ pkg-check
- ✓ pkg-install (ubuntu-20.04, 3.7)
- ✓ pkg-install (macOS-10.15, 3.7)
- ✓ pkg-install (windows-2019, 3.7)

Test combination of
- Hardware setup
- Operating system
- Python version
- Dependencies

Runs tests, docs, coverage, lintint, package install ect

Each PR triggers all tests



✓ Changes approved

2 approving reviews by reviewers with write access. Learn more.

Show all reviewers

✓ 2 approvals

3 pending reviewers

Some checks were not successful

35 successful, 1 cancelled, and 1 failing checks

Hide all checks

✓	CI testing - base / doctest (ubuntu-20.04, 3.7) (pull_request) Successful in 43s	Required	Details
✓	CI testing - complete / pytest (ubuntu-20.04, 3.6, minimal) (pull_request) Successful in 20m	Required	Details
✓	Check Code formatting / flake8 (pull_request) Successful in 11s	Required	Details
✓	Docs check / test-docs (pull_request) Successful in 1m	Required	Details
✓	Install pkg / pkg-check (pull_request) Successful in 15s	Required	Details

Required statuses must pass before merging

All required statuses and check runs on this pull request must run successfully to enable automatic merging.

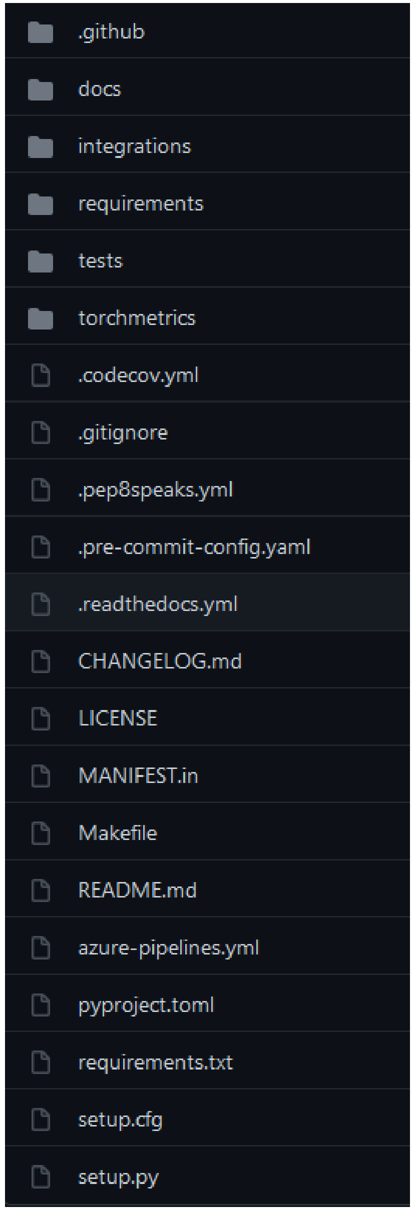
Auto-merge enabled

This pull request will squash and merge automatically when all requirements are met. Learn more

Disable auto-merge

Other files to consider

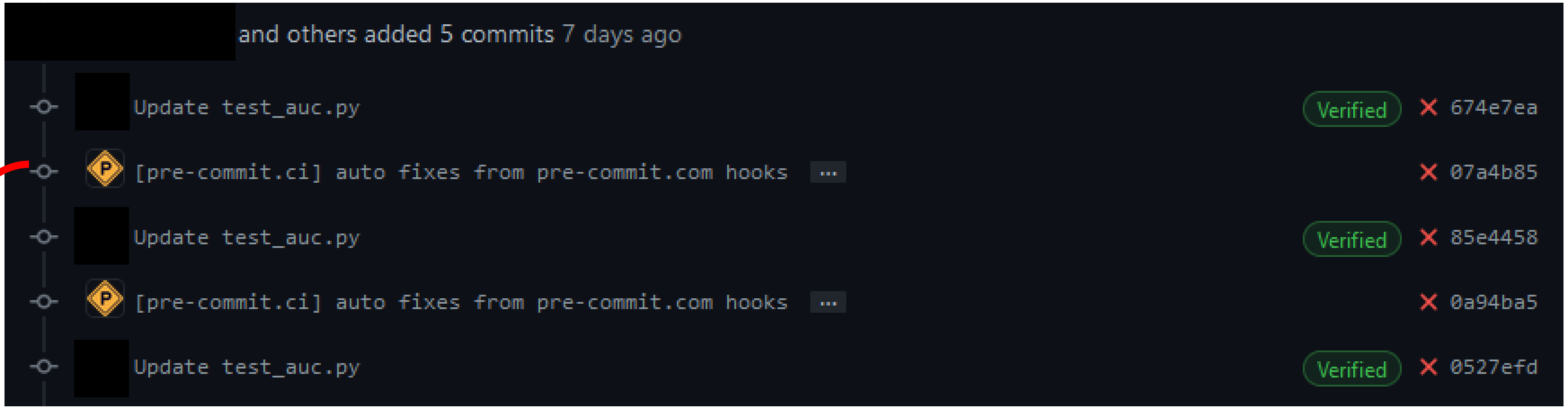
- reguirements.txt
- LICENSE
- README.md
- .gitignore



.github
docs
integrations
requirements
tests
torchmetrics
.codecov.yml
.gitignore
.pep8speaks.yml
.pre-commit-config.yaml
.readthedocs.yml
CHANGELOG.md
LICENSE
MANIFEST.in
Makefile
README.md
azure-pipelines.yml
pyproject.toml
requirements.txt
setup.cfg
setup.py

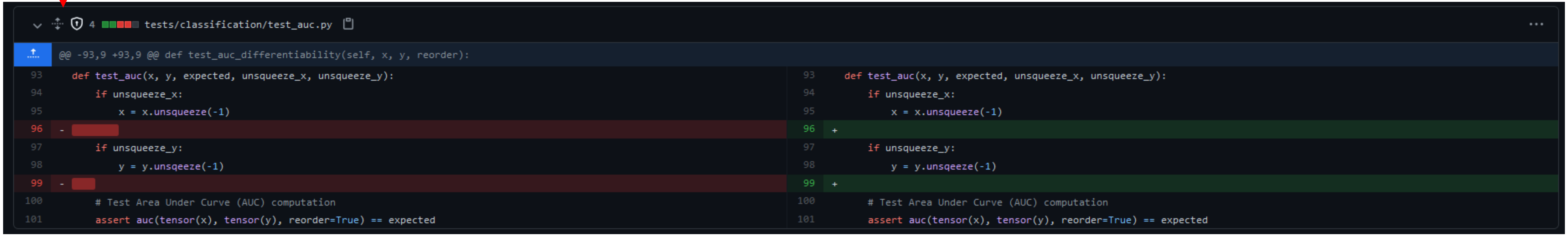
Advanced: Use bots

Bots can take care of tidious task for you (like linting)



[redacted] and others added 5 commits 7 days ago

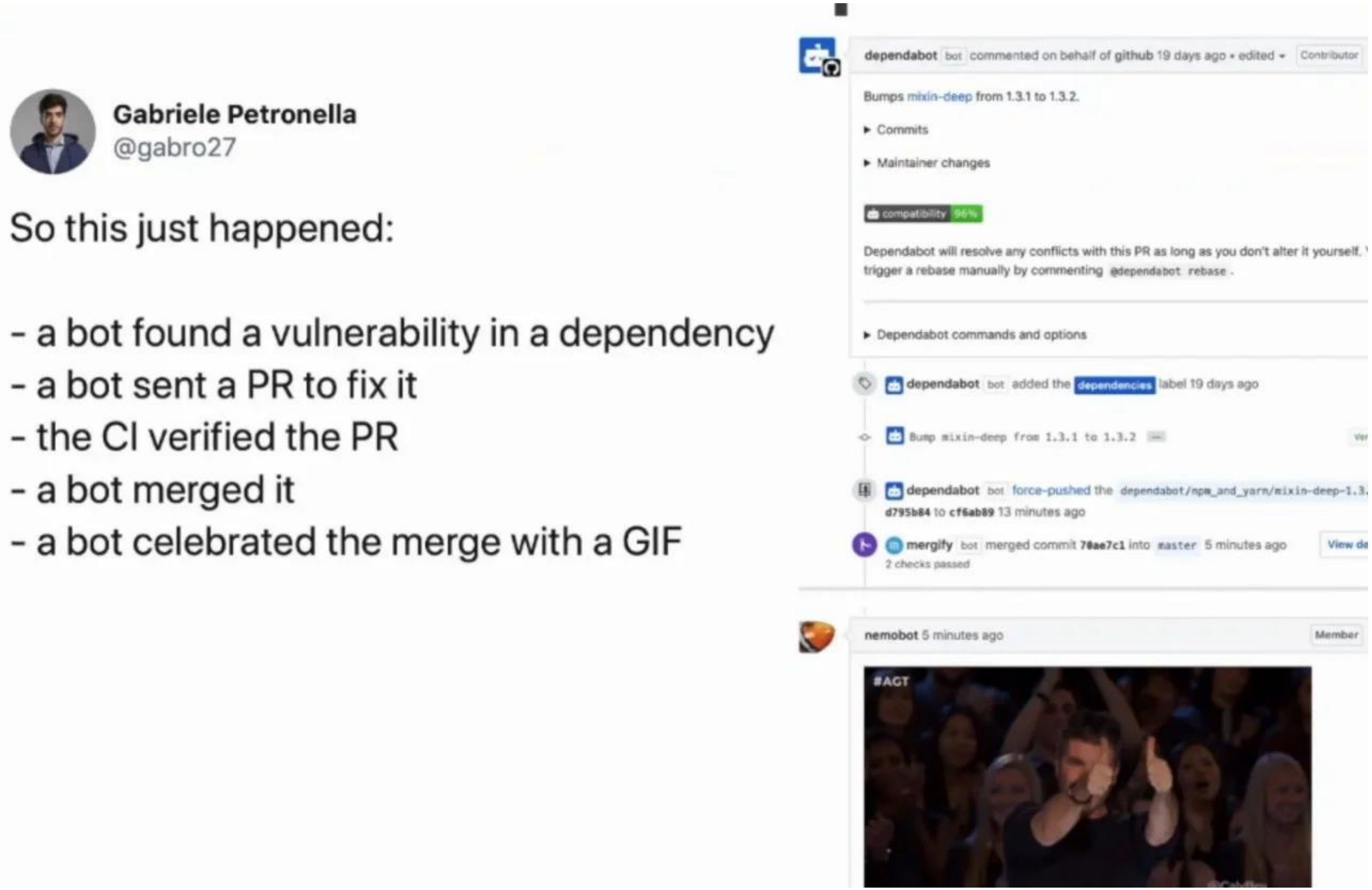
[redacted] Update test_auc.py	Verified	✗	674e7ea
[pre-commit.ci] auto fixes from pre-commit.com hooks …		✗	07a4b85
[redacted] Update test_auc.py	Verified	✗	85e4458
[pre-commit.ci] auto fixes from pre-commit.com hooks …		✗	0a94ba5
[redacted] Update test_auc.py	Verified	✗	0527efd



4 🟩🟩🟥🟥⬜ tests/classification/test_auc.py

```
@@ -93,9 +93,9 @@ def test_auc_differentiability(self, x, y, reorder):
 def test_auc(x, y, expected, unsqueeze_x, unsqueeze_y):      | def test_auc(x, y, expected, unsqueeze_x, unsqueeze_y):
     if unsqueeze_x:                                          |     if unsqueeze_x:
         x = x.unsqueeze(-1)                                  |         x = x.unsqueeze(-1)
-                                                             | +
     if unsqueeze_y:                                          |     if unsqueeze_y:
         y = y.unsqeeze(-1)                                   |         y = y.unsqeeze(-1)
-                                                             | +
     # Test Area Under Curve (AUC) computation                |     # Test Area Under Curve (AUC) computation
     assert auc(tensor(x), tensor(y), reorder=True) == expected |     assert auc(tensor(x), tensor(y), reorder=True) == expected
```

The future is here



Gabriele Petronella
@gabro27

So this just happened:

- a bot found a vulnerability in a dependency
- a bot sent a PR to fix it
- the CI verified the PR
- a bot merged it
- a bot celebrated the merge with a GIF

dependabot bot commented on behalf of **github** 19 days ago • edited ▾ Contributor

Bumps mixin-deep from 1.3.1 to 1.3.2.

▸ Commits

▸ Maintainer changes

compatibility 96%

Dependabot will resolve any conflicts with this PR as long as you don't alter it yourself. trigger a rebase manually by commenting `@dependabot rebase`.

▸ Dependabot commands and options

dependabot bot added the dependencies label 19 days ago

`Bump mixin-deep from 1.3.1 to 1.3.2` …

dependabot bot force-pushed the `dependabot/npm_and_yarn/mixin-deep-1.3.`
d795b84 to **cf6ab89** 13 minutes ago

mergify bot merged commit **70ae7c1** into `master` 5 minutes ago
2 checks passed

nemobot 5 minutes ago Member