

scikit-learn mentored internship

The scikit-learn team is opening a mentored and paid position for one person to work on the project for 5 months.

This would be a full time position, with regular in person sessions with a mentor, and virtual contact with mentors as needed. For this position, your mentor in Berlin would be Adrin Jalali (<https://github.com/adrinjalali/>), and Guillaume Lemaitre (<https://github.com/glemaitre/>) would be your second albeit remote mentor.

This mentored internship is part of the NumFocus diversity program and therefore we will only consider applications of people who identify as belonging to an underrepresented group in tech.

Location

Berlin. You'll need to be allowed to work and be paid in Germany.

Job Description

This position's goal is to familiarise you with the project's internals so that you can have deeper code contributions to the project. Ideally you would keep contributing after the internship, and one day you could be one of the core contributors of the project. We would start with some good first issues, documentation fixes, and easy code fixes, and work our way up to more complex topics.

One potential project for this position is to advance the work done for metadata routing (<https://scikit-learn-enhancement-proposals.readthedocs.io/en/latest/slep006/proposal.html>). That includes both documentation and implementation of the required tools and changes to the scikit-learn estimators. We acknowledge that's a complicated topic, and we'll make sure you have the tools and the knowledge to contribute to it. You are not expected to be able to sit and work on it alone.

Requirements

For you to be successful in this internship, the following skills are needed:

- Be able to work with git and GitHub, and follow a git/GitHub workflow
- Be comfortable with your IDE, whichever it might be: VSCode, PyCharm, vim, emacs, etc. This position requires more than scripting in Jupyter notebooks.
- Python: you don't need to be very proficient in Python, but knowing your way around, navigating and understanding documentation, understanding object oriented programming and software engineering skills would help.

- Communication: We work collaboratively on the project, and your pull requests will be reviewed by people other than your mentor. It's important to be able to communicate clearly and calmly to have a constructive flow on the pull request.
- Statistical machine learning: scikit-learn is a statistical machine learning library. You don't need to deeply understand all the algorithms in the library, but having a good grasp of machine learning algorithms helps with understanding the code base.

How We Work

We work by usually picking up issues from the issue tracker (<https://github.com/scikit-learn/scikit-learn/issues>), talking about the issue if certain things are not clear on the issue itself, and then submitting a pull request. You will choose the issues with your mentor since finding issues for people who are new to the project is always challenging.

There will almost always be a discussion on the pull request, and different people will raise their opinions. The author of the pull request is responsible to track all those discussions and apply corresponding suggestions to the pull request. You will work closely with your mentor to make sure you can understand and apply the feedback on your contributions.

For a more detailed explanation of what contributing to the project involves, you can have a look at our docs (<https://scikit-learn.org/dev/developers/contributing.html>). We know that's a very long document and we'll work with you to go through different sections of it.

We plan to have about 1 day worth of in person coworking with your mentor per week, and more virtual access to them as needed.

Payment

This position pays 2000€/month (brutto / gross), for 5 months. The money is paid as a stipend for the internship, and you'll need to handle the deductions such as health and social insurance yourself.

The payments will come from NumFocus (<https://numfocus.org/>) which is an NGO where scikit-learn is a financial affiliate. It has a European entity and you will be paid by that entity. You will need to handle your taxes and income declarations yourself.

How to Apply

You can apply by filling out this form: <https://forms.gle/HVyyeiLmER6A9ZaX9>