

WASP Assignment 1

REFLECTION ON PYNBLINT

UMEÅ UNIVERSITET

Minal Suresh Patil

My current research is focussed on machine explainability for multi-agent systems and computational argumentation. Before I started my PhD, I was working for a geospatial analytics startup where the use of `jupyter notebooks` was extensive and not as much at the moment in my PhD.

- What did you learn, in your individual session, about static analysis for ML and the `pynblint` tool?

The session was really useful and it helped me explore the various features of `notebooks` on a deeper level. The best practices were an important aspect of the session since during rapid prototyping I usually tend to overlook this particular aspect. Moreover, in a bigger setting such as working with teams or collaborators it is essential to use *Pynblint's* best practices.

- Will `pynblint` be useful to you in your WASP PhD project? Why or why not?

Pynblint will be useful when my research involves collaborating with teams and the requires large scale data analysis and issue of reproducibility has become a somewhat of a mandatory criterion in many ML conferences recently. Unfortunately, my team prefers R over other programming languages.

- Ideas for how the tool could be improved?

There were a couple of points we had discussed in order to make it better. First, introducing rules of linting that organisation(s) adhere to maybe in the form of formal training because the tool itself can be well understand and easy to use but to ensure consistency amongst different research groups in the same department. Second, the tool could be extended to R programming since almost all of them in my research group use R for the their research.

- What do you see as the limits for static analysis tools in ML? For code, models, and for data?

First, the issue false positive could easily arise in the context of ML. Sometimes, it is safe to ignore certain warnings and focus on the task at hand. Second, as the code-base increases in size I assume it does take a lot of time to analysis code and various models that is being developed.