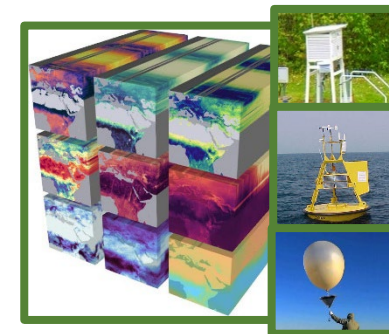ML@NFDI4Earth

Set Up of an Interest Group

Machine Learning at NFDI4Earth

Christopher Kadow – DKRZ

Christopher Irrgang - GFZ

et al.

Earth Science & Machine Learning & Research Data

Motivation

- Machine Learning is a rising branch in Earth System Science

- Big data: streams are needed, but need to be organized

- Scalability: HPC/clouds getting more and more into focus, however…

- ESS is a conservative research area in principle.
  - development cycles for methods, models, instruments, etc., often focus on the long term

    leads to data, hardware, methodic gaps within the community

- Gap to **other disciplines** developing ML in software and data is even bigger

Aim and Scope

IG ML@NFDI4Earth

is to maximize synergies and **bundle information** about ML activities in the field of ESS and to improve exchange between stakeholders across disciplinary boundaries.

will **gather and identify** projects, research topics, used techniques, best practices, software, frameworks, benchmarks, data sets, hardware setups, etc.

…

Concrete questions are:

- *How are ML frameworks installed at RDIs?*

- *How to interface common ML frameworks with data stored in the ESS?*

- *Have training, validation, test and verification data sets enough meta data?*

- *Do data management plans consider ML data sets already?*

- *Scalability of approaches from laptop to HPC?*

- *Are the interfaces available to fully exploit new ML in rather old Earth system modeling and analysis frameworks? (e.g. python / fortran / Julia interfaces)*

NFDI4Earth
Initial Member

ML@NFDI4Earth

Broad topics with Earth Science

Cross-connection to
NFDI4Earth-Academy

Cross-linking to
international ML communities

Hub strategy will be set up, not re-inventing the wheel, but focussing on the German **NFDI community**

Christopher Kadow - DKRZ, Hamburg
Christopher Irrgang - GFZ, Potsdam

kadow@dkrz.de

irrgang@gfz-potsdam.de