HELMHOLTZ AI | ARTIFICIAL INTELLIGENCE COOPERATION UNIT

DKRZ DEUTSCHES KLIMARECHENZENTRUM

Unlocking the Potential of ML for Earth and Environment Researchers

Frauke Albrecht, Caroline Arnold, Danu Caus, Harsh Grover, Andrey Vlasenko, Tobias Weigel

AI Consultants Earth & Environment @DKRZ

Helmholtz AI

Artificial Intelligence Cooperation Unit

Mission

Bring applied AI / ML techniques to your research questions and datasets

6 LOCAL UNITS

- Earth @ HZG
- Energy @ KIT
- Matter @ HZDR
- Space @ DLR
- Health @ HMGU
- Info @ FZJ

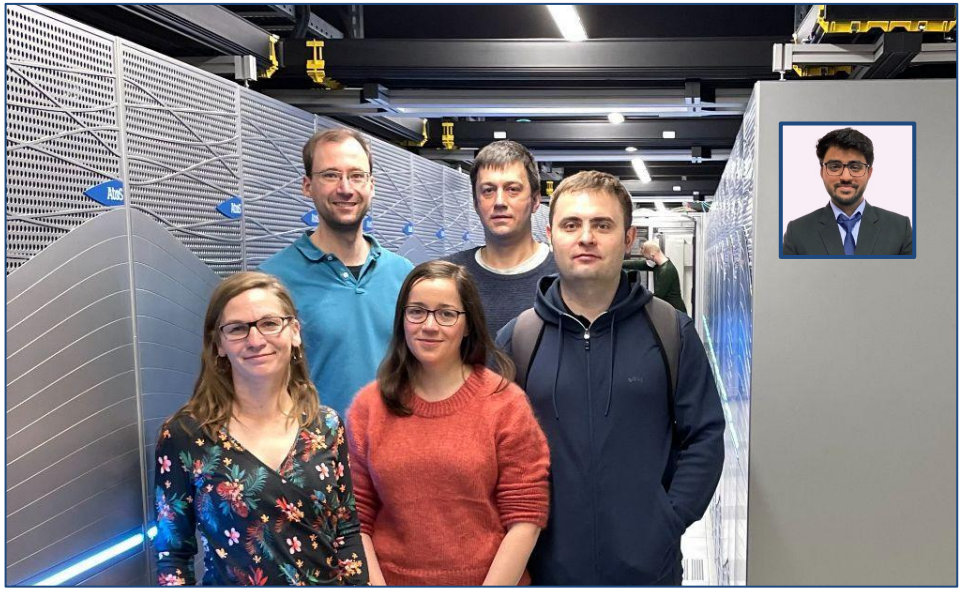
Each Unit

- Young Investigator Group
- AI Consultants

DKRZ DEUTSCHES KLIMARECHENZENTRUM

AI Consultants for Earth & Environment

DKRZ
DEUTSCHES KLIMARECHENZENTRUM



Vouchers from all
Earth & Environment
Helmholtz Centers

AWI
GFZ Helmholtz-Zentrum Potsdam
GEOMAR
UFZ HELMHOLTZ Zentrum für Umweltforschung
DLR Deutsches Zentrum für Luft- und Raumfahrt
JÜLICH Forschungszentrum
Helmholtz-Zentrum hereon
KIT

Vouchers

Selected user questions

Consultation

„I already use ML for my research project, however most of my colleagues work with different techniques.“

We act as a „sparring partner“, give feedback and suggestions, research minor issues, …

Implementation

„I have data at hand and I know what I want to achieve, but I need help with my ML research project.“

We implement code for ML training, data processing, …

Performance

„I have running ML code but it does not scale well / does not run fast enough.“

We support you with performance analysis and code review

Your question?

No voucher is like the other – contact us for help with your machine learning project!

Focus Areas for Support

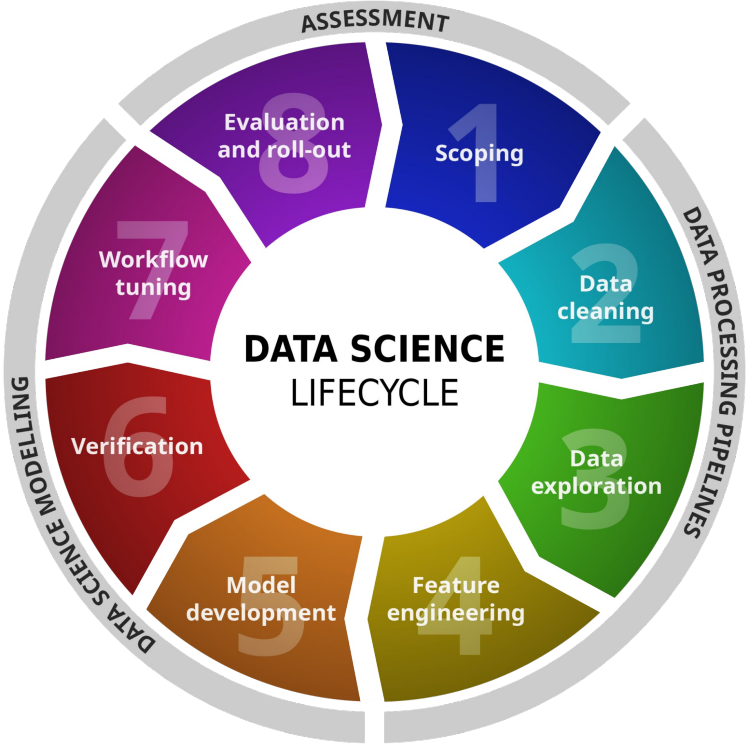
ML for Earth System Modelling

- **Building ESM-ML hybrids**
- Parametrizations for ICON
- Prototype ML models and integration with ESM code
- Interactive digital twins for extrapolation

ML on Earth observation data

- **Support for training, testing, tuning, deploying EO workflows with ML**
- Time series, example: Seismology
- Image-like data (e.g. jpg-images or Satellite images), example: Dead Sea Lakebed Segmentation

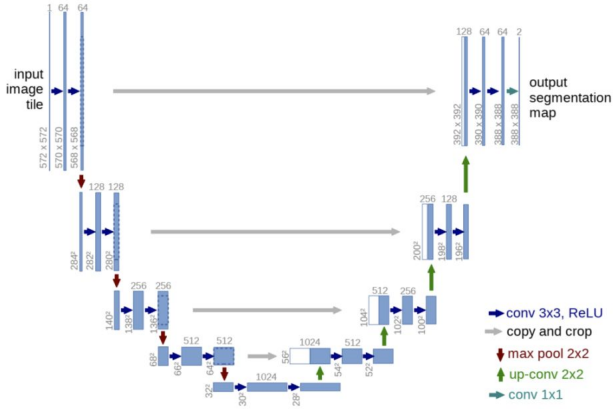
Full Data Science Lifecycle



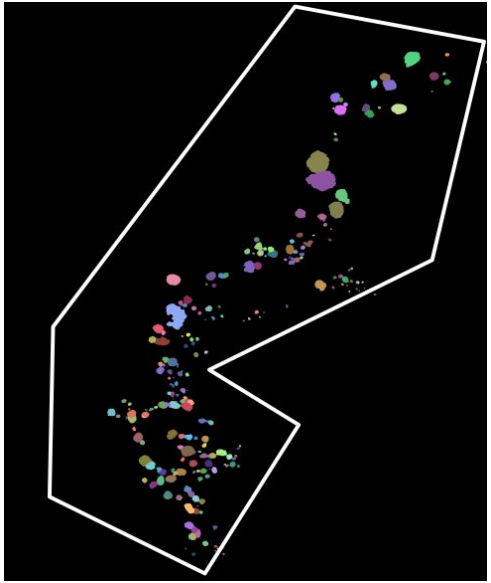
Dead Sea Lakebed Segmentation

Objective:
Perform sinkhole **Instance Segmentation** using deep learning techniques

Input



UNet Architecture



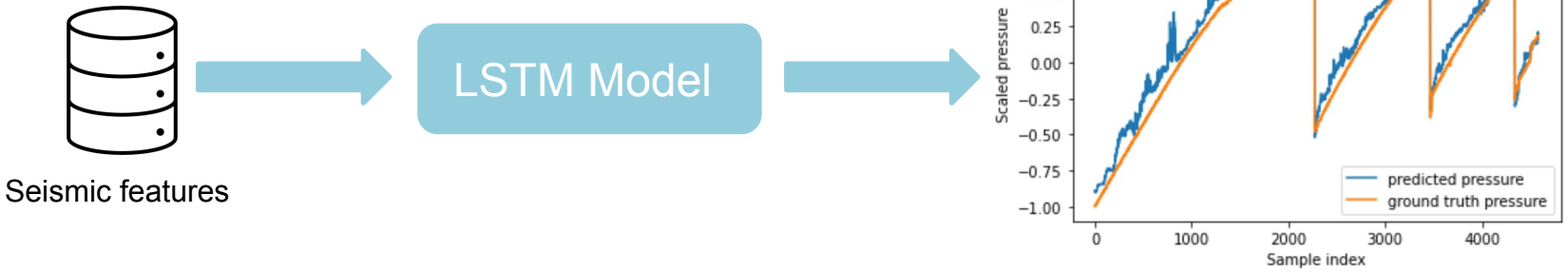
Sinkhole Instance Segmentation

Seismology

GFZ Helmholtz-Zentrum Potsdam

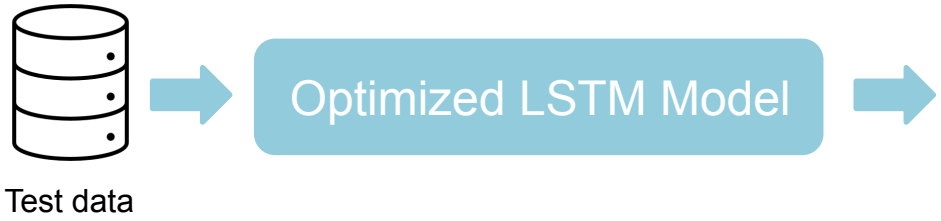
Objective:
Importance of **scientifically established** seismological features through the eyes of a neural network

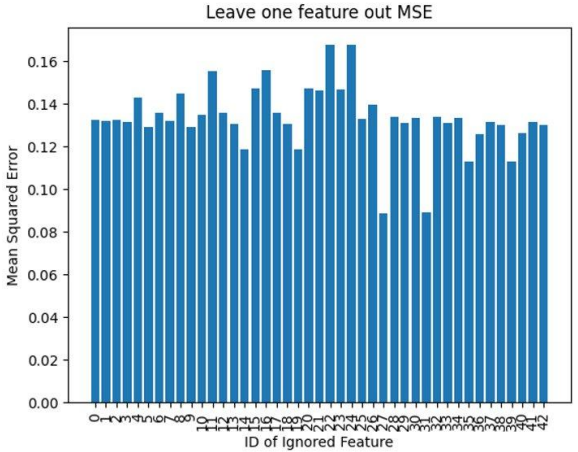
Phase 1:
Train optimal sequence model using a **Long Short Term Memory** neural network to predict pressure



Seismology

Phase 2:
Using **test** data and the **optimized** model from phase 1 to compute the error when one feature is ignored

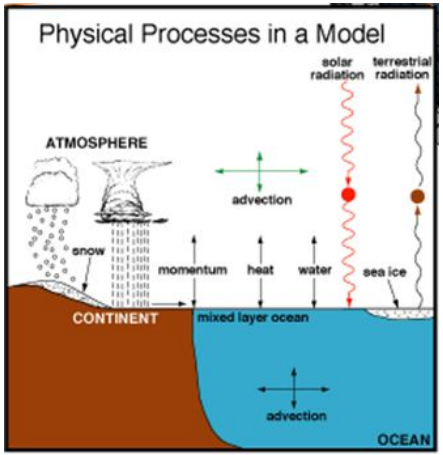




ML Emulators in Earth System Models (ESMs)

- ESM: few km grid resolution
- Processes on smaller scales, e.g.
  - convection
  - atmospheric chemistry
- Treatment in the ESM
  - parameterization
  - neglected due to computational effort

ICON: ICOsahedral Nonhydrostatic model



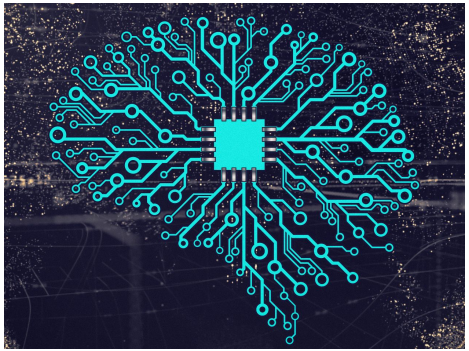
Replace sub-grid scale process by ML prediction

- "Easy" to train offline with good accuracy
- Need to test online - coupled to the ESM

Hybrid Models: Rapid Development

FORTRAN / CPU

Python / GPU or CPU



Voucher request

- Call Python ML inference inside Fortran ESM
- Should allow for iterative development - not hard coded
- Should be performant to run experiments

What is required for a good Voucher Problem?

(Big) data

A challenge

A client

Contact us to jumpstart your AI project:
consultant-helmholtz.ai@dkrz.de
Tobias Weigel, weigel@dkrz.de

https://www.helmholtz.ai
https://docs.dkrz.de/doc/software%26services/machine-learning/index.html