

# Python and scikit-learn for HPC

(PRACE Training Course)

February 25, 2021

VSB - Technical University of Ostrava

ONLINE

**Lecturers:** Stanislav Böhm, Georg Zitzlsberger - IT4Innovations

Data Science, Machine Learning, and Big Data are predominantly driven by Python. This one day course provides an insight into the most common Python extensions used to manage large data sets with Pandas, run distributed computations on HPC clusters with Dask, and leverage classical Machine Learning with scikit-learn.

Thursday February 25, 2021

08:45 - 09:00 Presentation  
09:00 - 10:30 Introduction to Pandas  
10:30 - 10:45 Coffee break  
10:45 - 12:00 Python in Distributed Environments  
12:00 - 13:00 Lunch break  
13:00 - 14:45 Introduction to scikit-learn  
14:45 - 15:15 Optimizations for CPUs, GPUs and numerical stability  
15:15 - 15:30 Coffee break  
15:30 - 16:45 Hands-on exercises  
16:45 - 17:00 Q&A



More information & registration  
[events.it4i.cz/event/113/](https://events.it4i.cz/event/113/)



This event is partially supported by The Ministry of Education, Youth and Sports from the Large Infrastructures for Research, Experimental Development and Innovations project "e-Infrastruktura CZ - LM2018140", partially by the PRACE-6IP project - the European Union's Horizon 2020 research and innovation programme under grant agreement No. 823767 and by the ExaQute project - the European Union's Horizon 2020 research and innovation programme under grant agreement No. 800898.