Moritz Boos

moritz.boos@gmail.com

ABOUT ME

I am a full-stack data scientist with over 8 years of experience in developing software and utilizing machine learning and statistics. I work end-to-end: From coordinating with stakeholders, data engineering, model development and deployment, to interpretation and communication of results. Before becoming a data scientist, I worked as a researcher at the intersection of machine learning and neuroscience.

SKILLS

General - Data science, research, software engineering, data engineering, communication

Programming - Python, SQL

Statistics - Bayesian & frequentist statistics

Machine Learning - Deep learning, (non-)linear regression and classification, unsupervised learning

Software - NumPy, Pandas, Scikit-Learn, Pytorch, Airflow, Docker, Google Cloud Project

EDUCATION

Ph.D. student in Computational Cognitive Neuroscience
University of Oldenburg
2016 - 2021

M.Sc. in Neurocognitive Psychology
University of Oldenburg
2013 - 2016

B.Sc. in Psychology
Technical University of Brunswick
2009 - 2013

LINKS

github
linkedin
website

EXTRAS

Mentoring - I mentor junior data scientists at the University of Sussex.

Academic exchange grant (2019) - I won a grant for developing a deep learning model for automatic speech recognition that is aligned to neuronal data Columbia University.

Communication - I presented my work at multiple conferences & regularly present projects within my company, as well as outside it.

EXPERIENCE

eyeo - Data Scientist 03/2021–
- I led a project that developed and productionized a model that increased the number of identified problematic ads by over 60%. Inferences of this model affect up to 250 million users.
- I designed and built data pipelines to ingest and process ad labeling data, while monitoring data quality (using Airflow, Docker, & GCP). This reduced manual work and flagged data issues that would have gone unnoticed.
- I built a statistical model for predicting user behavior and used it to consult stakeholders in key strategic decisions.
- Established best practices (e.g., Python development, code review, documentation) within the team.

University of Oldenburg - Researcher 11/2016–2/2021
- I conducted research and published articles on the neural basis of speech processing using machine learning and human neural data.
- I built research software (using Python and Docker) that was used by multiple people for machine learning in neuroimaging.
- I mentored junior researchers in statistics, machine learning, and Python programming.

Columbia University, New York - Visiting Researcher 6/2019–1/2020
- I developed a Deep Learning model for linking audio data with brain activity based on self attention in recurrent neuronal networks.
- I built data pipelines and machine learning experiment tracking tools for processing, cleaning, and linking audio data with human neural data.

University of Oldenburg - Part-time Research Data Scientist 2/2014–5/2015
- I developed models and built software for analyzing non-linearities in human invasive neural recordings.

University of Göttingen - Part-time Quantitative Researcher 2/2015–11/2015
- Statistical analysis of questionnaire data using R

Medical School Hanover - Part-time Research Data Scientist 11/2013–12/2015
- Statistical analysis of neural data using Python

SELECTED PUBLICATIONS

Journal Article 2021
Generalizable dimensions of human cortical auditory processing of speech in natural soundscapes: A data-driven ultra high field fMRI approach., Moritz Boos, Jörg Lücke, and Jochem W. Rieger., *In: NeuroImage (2021), p. 118106.*.

Preprint 2020
The role of auxiliary parameters in evaluating voxel-wise encoding models for 3T and 7T BOLD fMRI data., Moritz Boos, J. Swaroop Guntupalli, Jochem W. Rieger, and Michael Hanke., *In: biorxiv (2020). doi: 10.1101/2020.04.07.029397.*.

Journal Article 2016
Probabilistic inference: Task dependency and individual differences of probability weighting revealed by hierarchical Bayesian modelling., Moritz Boos, Caroline Seer, Florian Lange, and Bruno Kopp., *In: Frontiers in Psychology 7 (2016), p. 755.*.

OPEN SOURCE

Deep learning for auditory encoding 2022
- A Python library for training and interpreting deep recurrent models with self-attention for computational neuroscience.

Nilearn 2016–2020
- Multiple contributions to a Python machine learning library for neuroimaging.

Voxel-wise encoding 2020
- A Python library for running scalable encoding models on large neuroimaging datasets with Docker.