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 was 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 incorporates neuronal data.

Communication - I presented my work at multiple conferences.

EXPERIENCE

eyeo (Adblock Plus) - Senior Data Scientist 6/2022–

- Developed KPIs, identified opportunities for data science and data engineering to provide value and implemented solutions towards that end.
- Took strategic decisions to maintain a healthy and scalable data science tech stack by managing the integration of new tools and services within the existing infrastructure, leveraging Google Cloud Platform across multiple projects.
- Led the design and implementation of a data engineering system using Airflow, Docker, and GCP, thereby reducing manual tasks, enhancing data quality, and unblocking multiple downstream projects.
- Up-skilled team members in the use of modern software and data engineering practices.

- Data Scientist 03/2021-5/2022

- 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.
- Built a statistical model for predicting user behavior and used it to consult stakeholders in key strategic decisions.
- Established best software engineering practices within the team.

University of Oldenburg - Researcher 11/2016–2/2021

- Conducted research and published articles on the neural basis of speech processing using machine learning and human neural data.
- Built research software (using Python and Docker) for machine learning in neuroimaging.
- 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.

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.

Self-employed - Research Data Scientist 11/2013–11/2015

- Implemented a Bayesian statistical analysis of neural data using Python (for Medical School Hanover).
- Statistical analysis of questionnaire data using R (for University of Göttingen).

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.*.

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.

Voxel-wise encoding 2020

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

Open Source contributions

- Nilearn - machine learning for neuroimaging
- Skrub - machine learning with dirty categorical data