Clara Hoffmann

LinkedIn | Github | Stack Overflow

Passionate machine learner aiming to make deep learning more certain in its predictions by leveraging Bayesian statistics. I'm interested in developing scalable and safe solutions for deep learning and CV with efficient code and SOTA methodology. Currently, I am interested in weakly-/self-supervised learning and processing of complex inputs

Education

Ph.D. candidate in Computer Science - Center for Trustworthy Data Science and Security starting in 02/2023
Collaboration with Charité hospital, Berlin *Dortmund, Germany*
Calibrating deep predictive densities for modeling disease severity trajectories from MRI scans using copula-based on deep distributional regression

M.Sc. in Statistics - Humboldt University 10/2018 – 07/2021
Bayesian statistics, machine and deep learning. GPA: 1.2/1 *Berlin, Germany*
Master thesis: Marginally calibrated response densities for end-to-end (EtE) learning in autonomous driving (1.0)

B.Sc. in Economics - Humboldt University 10/2013 – 02/2017
Semester abroad at Maastricht University, Netherlands. GPA: 1.7/1 *Berlin, Germany*

Academic & Industry Experience

Machine Learning Engineer - Computer Vision 11/2021 – 12/2022
LiveEO - Satellite Based Infrastructure Monitoring *Berlin, Germany*
ML-based change detection for small objects using SAR satellite imagery, prototyping models and improved preprocessing, setting up scalable imagery workflows

Student Research Assistant - Bayesian Statistics 06/2021 – 10/2021
Chair of Applied Statistics - Humboldt University *Berlin, Germany*
Research project on copula-based marginally calibrated regression for discrete responses with spatially structured selection priors for Prof. Dr. Nadja Klein, deriving the model structure and implementing the model in code

Junior Data Scientist 12/2019 – 08/2020
idalab - Agency for Data Science *Berlin, Germany*
Developing and implementing ML-based solutions in health care and life sciences. Amongst others: deep automatic text summarization for scientific publications and scalable Lasso for multi-class classification

Intern for Statistical Consulting 09/2018 – 12/2019
fu:stat Statistical Consulting - Freie Universität Berlin *Berlin, Germany*
Consulting Master and doctoral students regarding statistical aspects of their research within the university-run statistical consulting department fu:stat

Student Research Assistant 09/2018 – 12/2019
Institute of Economic Policy - Humboldt University *Berlin, Germany*
Assisting with macroeconomic research about Dynamic General Equilibrium models in Matlab and Julia

Intern & Student Research Assistant 09/2017 – 09/2018
German Institute for Economic Research - Education & Family, German Socioeconomic Panel *Berlin, Germany*
Research on Germany's largest panel data set (SOEP), implementing regression for current research projects

Publications & Research Presentations

Marginally calibrated response distributions for EtE-learning in autonomous driving | *Arxiv preprint* 10/2021
Scalable estimation for marginally calibrated response densities to quantify the uncertainty of steering angles in deep end-to-end learners. Authors: Clara Hoffmann & Nadja Klein, accepted to Annals of Applied Statistics, Arxiv preprint available here

Marginally Calibrated Response Densities for EtE Learning | *Talk at Statistical Week 2021* 09/2021
Research presentation about scalable and reliable uncertainty quantification of end-to-end learners (manim-animated slides available here)

Awards

Prize of the German Statistical Association for the best Master thesis (DStatG) 09/2022

Technical Skills

Languages: Python (●●●●◐): NumPy, pandas, PyTorch, TensorFlow, Keras, Stan, joblib, Ray, rasterio, geopandas, unit testing, typing ; R (●●●●○); SQL (●●●○○), experience with handling GB/TB datasets, shell scripting
Tech: AWS (S3, EC2, Batch), Git, conda, Docker **Other**: Jira, Confluence, Scrum