

Kiril Klein, PhD

Data Scientist — Causal Inference

kiril.vadimovic.klein@gmail.com | +45 91 40 39 62 | Copenhagen, Denmark | open to relocation



PROFILE

Data Scientist and ML Engineer with a PhD in Machine Learning and Causal Inference. I specialize in estimating effects from observational data using target trial emulation, IPW, propensity scores, and multiple-testing methods. I own the full ML development cycle, from data requirements and model training to validation, containerization, and production integration.

CORE SKILLS

- **Causal inference and statistics:** target trial emulation, IPW, propensity scores, TMLE, DAGs, survival analysis, experimental design, and A/B testing.
- **Machine learning and data:** Python, PyTorch, SQL, Pandas, scikit-learn, XGBoost, transformers, NLP.
- **Engineering:** ETL pipelines, Docker, Git, Bash, CI/CD, Azure, HPC.
- **Languages:** English (professional); Russian and German (native); French and Danish (basic).

EXPERIENCE

Machine Learning Engineer

Aiomic, medical AI, Copenhagen | 09/2025 – present

- **ML architecture:** building the product ML stack from scratch, combining transformers with classical NLP/ML components (TF-IDF, XGBoost) in a single pipeline, from data preparation and training through production integration.
- **Validation:** leading blinded validation of the ML system on historical data from a large US clinical system, with a target cohort of approximately 16,000 patients.
- **Data requirements:** defining required data, identifying missing fields, and preparing technical data requests for the team working with the clinical partner.

PhD Fellow, PHAIR project (Innovation Fund Denmark)

Pioneer Centre for AI, DIKU, University of Copenhagen | 09/2022 – 03/2026

- Applied target trial emulation, IPW, and propensity scores to estimate treatment effects from Danish national registry data where randomized experiments are not feasible.
- Built an untargeted causal-screening pipeline to estimate effects across hundreds of outcomes with multiple-testing correction.
- Formalized hypotheses, constructed DAGs to control confounding, and compared effect-estimation methods across millions of EHR records.
- Coordinated development of PHAIR-EHR, a causal-inference extension of CORE-BEHRT, and made the largest individual contribution to the current collaborative [BONSAI](#) codebase in a five-person team.

Research Assistant

DIKU, University of Copenhagen | 09/2021 – 08/2022

- Built ETL pipelines for clinical metadata and brain MRI scans; maintained the dataset behind a series of ML projects and owned data quality and secure transfers.

SELECTED PUBLICATION

- **CORE-BEHT: A Carefully Optimized and Rigorously Evaluated BEHT**. *Proceedings of Machine Learning Research*, volume 252, 2024. (joint first authorship)

EDUCATION

PhD in Machine Learning and Causal Inference

University of Copenhagen | 09/2022 – 03/2026

- Dissertation: “*Scalable Causal Inference on Electronic Health Records Using Transformers*”.

MSc in Computational Physics

University of Copenhagen | 09/2020 – 06/2022

- Thesis: “*Brain Volumetric Changes due to COVID-19: A Retrospective Study on a Danish Cohort*” — retrospective MRI analysis using statistical and machine learning methods.

BSc in Physics

University of Erlangen-Nuremberg, Germany | 09/2017 – 08/2020

SPEAKING AND TEACHING

- **Invited speaker and panelist, D3A Conference, Denmark** (October 2024) — collaboration between domain experts and AI researchers.
- **Teaching Assistant, Advanced Topics in Deep Learning, Department of Computer Science, University of Copenhagen** (autumn 2024) — modern deep-learning architectures.
- **Teaching Assistant, ProbAI Summer School, Copenhagen** (June 2024) — sessions on causal inference and generative models for approximately 100 participants.