

Research Interests

Current: Scaling Laws for different model architectures. Large Language Models pretraining.
Past: Probabilistic and Simulation-based inference. Explainability.

Education

- 2023 – 2025 **University of Tübingen, Tübingen, Germany**
MSc *Machine Learning, Wübben Scholar, Distinction* Grade : 1.1/1
Master's thesis: AutoML for Simulation-based Inference supervised by Prof. Jakob Macke and Dr. Katharina Eggensperger.
- 2016 – 2020 **Indian Institute of Technology (IIT), Kharagpur, India**
Bachelor of Technology (B.Tech), Major in *Electrical Engineering* CGPA: 9.15/10
Minor in *Computer Science and Engineering* minor CGPA: 9.15/10

Research Positions

- December 2025 – current **ELLIS Institute Tübingen, Research Engineer, OpenEuroLLM**
I am working on different aspects of language models pretraining in the OpenEuroLLM project supervised by Dr. Aaron Klein and Dr. David Salinas.
- February 2025 – June 2025 **ELLIS Institute Tübingen, Student Research Assistant, with Dr. Wieland Brendel**
Sparse pretraining of dynamic Mixture-of-Experts (MoE) models based on the Tokenformer architecture.
Experimented with training MoE models with different tricks to avoid instabilities.
- January – July, 2024 **Universitätsklinikum Tübingen, Student Research Assistant, with Dr. Vadim Borisov**
Identifying influential species associated with colorectal cancer (CRC) using rule-based models and symbolic regression from microbiome data (in Dr. med. Christoph Stein-Thoeringer's group).
- August, 2020 – September, 2023 **IBM Research, Bangalore, India, Research Software Engineer**
Worked on problems in explainability (understanding model differences, certifying existing explanations), bias and individual fairness (in tabular and Speech-to-Text settings) and API testing.

Selected Research Projects

- May 2025 – October 2025 **AutoML for Simulation-based Inference (SBI), Master's thesis, with Dr. Jakob H. Macke, Dr. Katharina Eggensperger & Guy Moss**
Worked on automating the SBI pipeline involving different posterior inference methods and performance metrics for easy application on new inference tasks. [Poster](#) for the Amortized ProbML Workshop at EurIPS 2025. [\[Thesis\]](#)
- Sept. 2024 – Feb. 2025 **Sequential Source Density Estimation, Research project, with Guy Moss and Dr. Jakob H. Macke**
Implemented a sequential source distribution estimation method that is demonstrably 30-50% simulation efficient compared to the non-sequential variant. [\[Report\]](#) [\[Code\]](#)

Selected papers

- C = Conference, B = Book chapter A = arXiv preprint/under review [\[Google scholar\]](#)
- [A3] Abhash Kumar Jha*, Diana Alexandra Onuțu*, Neeratoy Mallik*, **Swagatam Haldar***, Sam Laing, Niccolò Ajroldi, Shiwei Liu, Joaquin Vanschoren, Aaron Klein. "Amortizing Scaling Law Construction Costs," 2026, *Equal contribution. [\[arXiv\]](#)

- [A2] **Swagatam Haldar**, Christoph Stein-Thoeringer, Vadim Borisov. "*Interpreting Microbiome Relative Abundance Data Using Symbolic Regression*," 2024. [\[arXiv\]](#) [\[Code\]](#)
- [C2] Amit Dhurandhar, **Swagatam Haldar**, Dennis Wei, Karthikeyan Natesan Ramamurthy. "*Trust Regions for Explanations via Black-Box Probabilistic Certification*," International Conference on Machine Learning, (ICML), July 2024. [\[Paper\]](#) [\[Code\]](#)
- [C1] **Swagatam Haldar**, Diptikalyan Saha, Dennis Wei, Rahul Nair, Elizabeth M. Daly. "*Interpretable Differencing of Machine Learning Models*," Conference on Uncertainty in Artificial Intelligence (UAI), August 2023. [\[Paper\]](#) [\[Code\]](#)
- [B1] Diptikalyan Saha, Aniya Agarwal, Sandeep Hans, **Swagatam Haldar**. "*Testing, Debugging, and Repairing Individual Discrimination in Machine Learning Models*," Ethics in Artificial Intelligence: Bias, Fairness and Beyond, December 2023. [\[Chapter\]](#)
- [A1] **Swagatam Haldar**, Deepak Vijaykeerthy, Diptikalyan Saha. "*Automated Testing of AI Models*," 2021. [\[arXiv\]](#)

Scholastic Achievements

- 2024 *Wübben Stiftung Wissenschaft* Student Grant for academic performance in master's studies.
- 2020 Selected for a research internship at the University of Alberta under the UARE program.
- 2016 Mamraj Agarwal Rashtriya Puraskar, by Governor of West Bengal, India.
- 2016 Secured State Rank *1st* in Class 12 (Higher Secondary) Examination, West Bengal, India.
- 2014 Secured State Rank *9th* in Class 10 (Secondary) Examination, West Bengal, India.

Other Activities

- Teaching assistant for the Mathematics for ML lecture by Prof. Dr. Ulrike von Luxburg, Winter 24/25.
- [Contributed](#) to Stack Overflow community with 2600+ reputation points, mostly before ChatGPT was a thing!
- Hobbies: Running, Bengali dramatics (at IIT Kharagpur), Learning magic tricks.