Maurice Weber

AI Researcher

mauriceweber.github.io | Maurice Weber | mauriceweber | maurice.weber@hotmail.com

Experience

Together AI | AI Researcher

January 2024 - present | Zurich, Switzerland

- Leading the design, construction, and release of large-scale datasets and driving the development of efficient LLMs for speculative decoding, covering the full development cycle including data processing, model training and evaluation.
- The resulting models and datasets form core components of Together AI's production inference platform and internal research efforts, demonstrably improving system capabilities.
- Led the development and release of the RedPajama datasets (100T+ tokens), which are among the most popular open LLM pre-training datasets and have been widely adopted in both industry and academia.

Xanadu Quantum Technologies | Research Intern

May 2022 - August 2022 | Toronto, Canada

- Co-first authored a paper on transformer-based generative models for quantum state tomography.
- Contributed to the development of the PennyLane quantum ML library and released cookbooks and tutorials.

Pictet Asset Management | Intern Multi-Asset Investments (part-time)

May 2016 - May 2018 | Zurich, Switzerland

- Assisted the multi asset teams in Geneva and Zurich in their daily investment process and developed tools for analysis of portfolio exposure to currencies, regions and industries.

UBS Investment Bank | Intern Equity Derivatives Trading

May 2015 - April 2016 | Zurich, Switzerland

- Assisted the derivatives trading teams in their daily routines.
- Developed trade supporting tools using MS Visual Basic, including a tool for real-time monitoring of competitor prices.

Education

ETH Zurich Ph.D. in Computer Science

February 2020 – December 2023 | Zurich, Switzerland

Exploring robustness guarantees for machine learning systems and quantum algorithms, in the context of adversarial attacks, naturally occurring noise and distribution shifts.
Supervisors: Prof. Ce Zhang & Prof. Martin Vechev

ETH Zurich MSc Mathematics (graduated with distinction)

September 2016 – May 2019 | Zurich, Switzerland

Coursework with focus on machine learning and mathematical statistics. Master thesis on lossy image compression with recurrent neural networks.
Supervisors: Prof. Ce Zhang & Prof. Nicolai Meinshausen.

GPA: 5.77 / 6.0

ETH Zurich BSc Mathematics

September 2011 – June 2015 | Zurich, Switzerland

Courses covered wide areas of Mathematics including Linear Algebra, Functional Analysis, Topology, Probability Theory, Brownian Motion & Stochastic Calculus. Bachelor thesis on value at risk with defaultable securities.

GPA: 4.85 / 6.0

Profile

AI Researcher specializing in Large Language Models (LLMs), large-scale data processing, and efficient architectures. Proven experience leading the full development pipeline for multi-billion parameter models and constructing foundational open-source datasets (RedPajama). PhD from ETH Zurich with publications in top-tier venues (NeurIPS, ICML, S&P, CCS) focused on trustworthy ML and verifiable robustness.

Skills

Deep Learning Frameworks PyTorch, PyTorch FSDP, DeepSpeed, HuggingFace Transformers, vllm.

Programming Languages Python, Bash, SQL, R, Mathematica.

Data Pipelines & Scale Ray, Daft, Polars, Pandas, NumPy.

Tools & Infrastructure Git, Docker, Kubernetes (k8s), AWS, wandb.

Expertise

- Large Language Models
- Machine Learning
- Scientific Research
- Software Development

Languages

German: ●●●●● Native
English: ●●●●○ Fluent
Spanish: ●●●●○ Fluent
French: ●●●○○ Intermediate

Service

Open Source Led the development of the RedPajama datasets, among the largest and most widely adopted open-source datasets for pretraining Large Language Models (4.5k GitHub stars, 2M Downloads). Drove project from conception to public release and community adoption.

Reviewer Served as a reviewer for NeurIPS, Physical Review A, Physical Review Research, PRX Quantum, npj Quantum Information, IEEE Transactions on Neural Networks and Learning Systems.

Select Publications

(* denotes equal contribution)

2024 **Maurice Weber**, Dan Y. Fu, Quentin Anthony, Yonatan Oren, Shane Adams, Anton Alexandrov, Xiaozhong Lyu, Huu Nguyen, Xiaozhe Yao, Virginia Adams, Ben Athiwaratkun, Rahul Chalamala, Kezhen Chen, Max Ryabinin, Tri Dao, Percy S Liang, Christopher Ré, Irina Rish, and Ce Zhang. Redpajama: an open dataset for training large language models. In Advances in Neural Information Processing Systems, 2024

2023 **Maurice Weber***, Carlo Siebenschuh*, Rory Marshall Butler*, Anton Alexandrov, Valdemar Ragnar Thanner, Georgios Tsolakis, Haris Jabbar, Ian Foster, Bo Li, Rick Stevens, and Ce Zhang. Wordscape: a pipeline to extract multilingual, visually rich documents with layout annotations from web crawl data. In Advances in Neural Information Processing Systems, 2023

2023 **Maurice Weber***, Xiaojun Xu*, Bojan Karlaš, Ce Zhang, and Bo Li. Rab: Provable robustness against backdoor attacks. In 2023 IEEE Symposium on Security and Privacy (SP), 2023

2022 Haoxiang Wang*, **Maurice Weber***, Josh Izaac, and Cedric Yen-Yu Lin. Predicting properties of quantum systems with conditional generative models. arXiv preprint arXiv:2211.16943, 2022

2022 **Maurice Weber**, Linyi Li, Boxin Wang, Zhikuan Zhao, Bo Li, and Ce Zhang. Certifying out-of-domain generalization for blackbox functions. In 39th International Conference on Machine Learning, 2022

2022 **Maurice Weber**, Abhinav Anand, Alba Cervera-Lierta, Jakob S Kottmann, Thi Ha Kyaw, Bo Li, Alán Aspuru-Guzik, Ce Zhang, and Zhikuan Zhao. Toward reliability in the nisq era: Robust interval guarantee for quantum measurements on approximate states. Physical Review Research, 4(3):033217, 2022

2022 Mintong Kang*, Linyi Li*, **Maurice Weber**, Yang Liu, Ce Zhang, and Bo Li. Certifying some distributional fairness with subpopulation decomposition. In Advances in Neural Information Processing Systems 35 (NeurIPS 2022), 2022

2021 **Maurice Weber**, Nana Liu, Bo Li, Ce Zhang, and Zhikuan Zhao. Optimal provable robustness of quantum classification via quantum hypothesis testing. npj Quantum Information, 7(1):1–12, 2021

2021 Linyi Li*, **Maurice Weber***, Xiaojun Xu, Luka Rimanic, Bhavya Kailkhura, Tao Xie, Ce Zhang, and Bo Li. Tss: Transformation-specific smoothing for robustness certification. In 2021 ACM SIGSAC Conference on Computer and Communications Security (CCS '21), 2021

2020 **Maurice Weber**, Cedric Renggli, Helmut Grabner, and Ce Zhang. Observer dependent lossy image compression. In 42nd German Conference on Pattern Recognition, 2020

2019 Filipe Barata, Kevin Kipfer, **Maurice Weber**, Peter Tinschert, Elgar Fleisch, and Tobias Kowatsch. Towards device-agnostic mobile cough detection with convolutional neural networks. In 2019 IEEE International Conference on Healthcare Informatics (ICHI), pages 1–11. IEEE, 2019

Conferences & Workshops

Dec 2024 **Advances in Neural Information Processing Systems 37, 2024** | Vancouver, B.C., Canada
Spotlight Poster: RedPajama: an Open Dataset for Training Large Language Models

Dec 2023 **Advances in Neural Information Processing Systems 36, 2023** | New Orleans, LA, US
Poster: WordScape: a Pipeline to extract multilingual, visually rich Documents with Layout Annotations from Web Crawl Data

May 2023 **IEEE Symposium on Security and Privacy (SP) 2023** | San Francisco CA, US
Talk: RAB: Provable Robustness against Backdoor Attacks

July 2022 **39th International Conference on Machine Learning (ICML) 2022** | Baltimore MD, US
Spotlight Talk: Certifying Out-of-Domain Generalization for Blackbox Functions

Mar 2022 **26th Conference on Quantum Information Processing (QIP) 2022** | Los Angeles CA, US
Poster: Toward Reliability in the NISQ Era: Robust Interval Guarantee for Quantum Measurements on Approximate States

Feb 2021 **25th Conference on Quantum Information Processing (QIP) 2021** | Virtual
Poster: Optimal Provable Robustness of Quantum Classification via Quantum Hypothesis Testing Talk: Optimal Provable Robustness of Quantum Classification via Quantum Hypothesis Testing

Sep 2020 **42nd German Conference on Pattern Recognition (GCPR) 2020** | Virtual
Poster: Provable Robust Learning Based on Transformation-Specific Smoothing

Teaching

Fall 2022 **Programming (C++) (D-MATH/D-PHYS ETH Zurich)** | Teaching Assistant

Spring 2022 **Data Modelling and Databases (D-INFK ETH Zurich)** | Teaching Assistant

Fall 2020 **Information Systems for Engineers (D-INFK ETH Zurich)** | Teaching Assistant

Spring 2017 **Linear Algebra II (D-MATH ETH Zurich)** | Teaching Assistant

Spring 2015 **Calculus II (D-MATH ETH Zurich)** | Teaching Assistant

References

Available upon request.