Maximilian Beck

PhD student with focus on efficient Large Language Model architectures.

Nationality German Languages German (native), English (fluent) Karlsruhe, Germany

maxbeck.ai maximilianbeck@live.de maxmbeck in/maxmbeck maximilianmbeck

Experience

Research Scientist Intern
Meta FAIR

May 2025 – Oct 2025 Paris, France

- Research on language models for code
- Evaluation of code execution prediction capabilities
- Exploration of architectures for code world models

PhD Researcher
NXAI GmbH

Aug 2024 – Ongoing Linz, Austria

- Research on scaling xLSTM as Transformer alternative
- Developing custom Triton and CUDA Kernels
- Training 7B parameter models with 100+ GPUs

Student Researcher
FZI Research Center for Information Technology

Nov 2019 – Jun 2021 Karlsruhe, Germany

- Research on autonomous driving and motion planning under uncertainty
- Software development for the internal driving simulator; contributed a package for visibility calculations with CGAL written in C++ and Python (github repo)

Student Research Assistant
wbk Institute for Production Science, KIT

Jan 2017 – Aug 2018 Karlsruhe, Germany

- PLC programming (tool change, pallet change, ...) for a Siemens CNC automation system
- Conception and design of a test rig for visual measuring of position accuracy in industrial robots

Education

PhD, Artificial Intelligence
Johannes Kepler University

Jul 2021 – Ongoing Linz, Austria

- PhD Advisor: Dr. Sepp Hochreiter
- First 1.5 years: Research on Few-Shot Learning, Domain Adaptation, and Loss Landscapes
- Since then: Research on efficient Large Language Model architectures

Master of Science, Mechatronics and Information Technology
Karlsruhe Institute of Technology - KIT

Oct 2017 – April 2021 Karlsruhe, Germany

- Focus: Control Technology in Mechatronics
- Master Thesis at Institute of Measurement and Control Systems (MRT) & FZI: *Motion Planning for Automated Vehicles in Uncertain Environments* (pdf); advised by Ömer Şahin Taş
- Final Grade: 1.4 (German grading system)

Exchange Program, Computer Engineering
San José State University

Aug 2018 – May 2019 San José, California, USA

- 2 semester exchange program. GPA: 3.8/4.0

Bachelor of Science, Mechatronics and Information Technology
Karlsruhe Institute of Technology - KIT

Oct 2014 – Oct 2017 Karlsruhe, Germany

- Bachelor Thesis at wbk Institute for Production Science: *Machine Control and Industry 4.0 Functions for a Machining Center of Type DMC 60H*
- Final Grade: 1.5 (German grading system)

Skills

- **Programming Languages:** Python (2018-present), C/C++ (2016-present), Java (2011-2016)
- **Other skills:** PyTorch (2021-present), JAX (2024-present), CUDA (2023-present), Triton (2024-present), slurm (2021-present), git (2015-present)

Selected Publications (Google Scholar)

- Maximilian Beck, Korbinian Pöppel, Phillip Lippe, Sepp Hochreiter: **Tiled Flash Linear Attention: More Efficient Linear RNN and xLSTM Kernels**. Arxiv, 2025. (link)
- Maximilian Beck*, Korbinian Pöppel*, Phillip Lippe*, Richard Kurle, Patrick M. Blies, Günter Klambauer, Sebastian Böck, Sepp Hochreiter: **xLSTM 7B: A Recurrent LLM for Fast and Efficient Inference**. International Conference on Machine Learning (ICML), 2025. (link)
- Maximilian Beck*, Korbinian Pöppel*, Markus Spanring, Andreas Auer, Oleksandra Prudnikova, Michael Kopp, Günter Klambauer, Johannes Brandstetter, Sepp Hochreiter: **xLSTM: Extended Long Short-Term Memory**. Advances in Neural Information Processing Systems (NeurIPS), 2024. (link) **[Spotlight]**
- Korbinian Pöppel, Maximilian Beck, Sepp Hochreiter: **FlashRNN: Optimizing Traditional RNNs on Modern Hardware**. International Conference on Learning Representations (ICLR), 2025. (link)
  - *Own contributions:* Contributed the FlashRNN sLSTM & LSTM Triton Kernels. Set up speed experiments. Contributed to writing.
- Benedikt Alkin, Maximilian Beck, Korbinian Pöppel, Sepp Hochreiter, Johannes Brandstetter: **Vision-LSTM: xLSTM as Generic Vision Backbone**. International Conference on Learning Representations (ICLR), 2025. (link)
  - *Own contributions:* Provided the mLSTM implementation.
- Thomas Schmied, Thomas Adler, Vihang Patil, Maximilian Beck, Korbinian Pöppel, Johannes Brandstetter, Günter Klambauer, Razvan Pascanu, Sepp Hochreiter: **A Large Recurrent Action Model: xLSTM enables Fast Inference for Robotics Tasks**. Arxiv, 2024. (link)
  - *Own contributions:* Contributed fast Triton inference kernels.
- Marius-Constantin Dinu, Markus Holzleitner, Maximilian Beck, Hoan Duc Nguyen, Andrea Huber, Hamid Eghbal-zadeh, Bernhard A. Moser, Sergei Pereverzyev, Sepp Hochreiter, Werner Zellinger: **Addressing Parameter Choice Issues in Unsupervised Domain Adaptation by Aggregation**. International Conference on Learning Representations (ICLR), 2023. (link) **[Oral]**
  - *Own contributions:* Implementation of our method IWA & baselines. Design of experiments and presentation of empirical results (see here). Contributed to writing.
- Martin Gauch, Maximilian Beck, Thomas Adler, Dmytro Kotsur, Stefan Fiel, Hamid Eghbal-zadeh, Johannes Brandstetter, Johannes Kofler, Markus Holzleitner, Werner Zellinger, Daniel Klotz, Sepp Hochreiter, Sebastian Lehner: **Few-Shot Learning by Dimensionality Reduction in Gradient Space**. Conference on Lifelong Learning Agents (CoLLAs), 2022. (link) **[Oral]**
  - *Own contributions:* Experiments and baseline implementations (see here). Made an animation for our method (see here). Contributed to writing.

* denotes equal contribution

I have served as a reviewer for the following conferences: NeurIPS-2024, ICLR-2025, ICML-2025, NeurIPS-2025.

Invited Talks

- Apr 2025 (London, UK): Invited Talk **xLSTM: New Architectures for Large Language Models** at G-Research
- Apr 2025 (Virtual): Presented **xLSTM 7B: A Recurrent LLM for Fast and Efficient Inference** online on ploutos.dev
- **xLSTM: Extended Long Short-Term Memory**
  - Jan 2025 (Virtual): Guest Lecture FAU Erlangen-Nuremberg in the *Machine Learning for Time Series* course
  - Nov 2024 (Munich, Germany): Opening Keynote at Machine Learning Week Europe
  - Sep 2024 (Virtual): Meta, Monetization AI Speaker Series
  - Sep 2024 (Karlsruhe, Germany): KIT, Autonomous Learning Robots (ALR) Lab
  - Aug 2024 (Vienna, Austria): ELLIS Seminar at Institute of Science and Technology Austria (ISTA)
  - Aug 2024 (Virtual): Google Research NLP Seminar
  - Jul 2024 (Vienna, Austria): Oral at ES-FoMo@ICML24
  - Jul 2024 (YouTube): 1littlecoder's youtube channel (Recording)
  - Jun 2024 (Helsinki, Finland): ELLIS PhD Spotlight Presentation at ELISE Wrap Up Conference 2024
  - Jun 2024 (Virtual): Intel Labs
- **Few-Shot Learning by Dimensionality Reduction in Gradient Space**
  - Oct 2022 (Virtual): Ruhr-University Bochum, Chair for Machine Learning