

Maximilian Beck

PhD student focusing on efficient sub-quadratic Large Language Models.

Nationality German Languages German (native), English (fluent)  Linz, Austria

 maxbeck.ai  [@maximilianbeck@live.de](mailto:maximilianbeck@live.de)  [maxmbeck](https://twitter.com/maxmbeck)  [in/maxmbeck](https://in.linkedin.com/in/maxmbeck)
 [maximilianmbeck](https://github.com/maximilianmbeck) [List of Publications](#)

EXPERIENCE

PhD Researcher

NXAI GmbH

 Aug 2024 – Ongoing  Linz, Austria

LLM Pretraining LLM Evaluation CUDA/Triton

- Contributed the [xLSTM](#) research project training and evaluation codebase, which is used internally by a team of 5+ researchers
- Working on scalability and efficiency of the xLSTM towards models with 7B+ parameters
- Writing custom Triton and CUDA Kernels
- Monitoring large training runs with 100+ GPUs

Student Researcher

FZI Research Center for Information Technology

 Nov 2019 – Jun 2021  Karlsruhe, Germany

Autonomous Driving Motion Planning C++

- Research on autonomous driving and motion planning under uncertainty; hosted by [Ömer Şahin Taş](#)
- Contributed a package for visibility calculations with CGAL to the internal driving simulator 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

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)

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 sub-quadratic Large Language Models

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)

PUBLICATIONS (LIST OF PUBLICATIONS)

- 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*: Worked on our method & baseline implementations, experiment design, 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*: Worked on experiments and baseline implementations ([see here](#)). Made an animation for our method ([see here](#)). Contributed to writing.

* denotes equal contribution

INVITED TALKS

- **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, German): 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](#)