Lukáš Dadák

Software Engineer · LLM Systems · GPU Performance

lukas.dadax@gmail.com · github.com/lukdmine

PROFILE

Software engineer and MSc student building LLM systems. Author of CACAO, an agentic CUDA kernel optimizer whose generated kernel outperformed the proposed llama.cpp kernel by up to 3×; building RAG features in production at Avenga. Background in GPU performance, systems programming and applied ML. Former competitive skiff sailor.

EXPERIENCE

Software Engineer

Aug 2025 – Present

Avenga

- Maintained and refactored a React/TypeScript frontend codebase, including CSS cleanup and library rewrites.
- Implemented RAG-based features and image-processing components.
- Worked across frontend and AI/ML-related functionality in a production software environment.

PROJECTS

CACAO — CUDA Agentic Coding Autotuning Optimizer

Bachelor thesis, June 2026 · ongoing research at FI MU · github.com/lukdmine/CACAO

- Agentic tree search where LLMs write CUDA kernels and their tuning spaces; validation, KTT autotuning and Nsight Compute profiling stay outside the model's control. Runs open-weight reasoning models unattended for up to 37 h.
- Generated a fused Gated DeltaNet kernel for llama.cpp: 2.7–3.0× faster than the kernel then proposed for llama.cpp and ahead of vLLM's Triton kernel on an RTX 3090 (+5.5 % end-to-end prefill on Qwen3.8-27B); design adopted by the maintainer's fused kernel.
- Scored 92–129 % of the best student kernel across six task/GPU cells of a GPU programming course (59–117 kernels each), ranking first on two, using an earlier version of the system and earlier open-weight models (Kimi k2.6, Qwen 3.5, GLM-5).

Battery recharging optimization with solar panels

Deployed in production

Modeled a battery charging strategy using machine learning and linear programming to optimize energy usage.

Neural network and CNN from scratch in Rust

Coursework · MUNI FI

Implemented neural network and convolutional network components without ML or linear algebra libraries.

EDUCATION

Masaryk University, Faculty of Informatics

MSc	Machine Learning and Artificial Intelligence	2026 – present
BSc	Programming and Application Development	2023 – 2026

SKILLS

Languages	Python, TypeScript, C, Java
LLM systems	agentic tool-use loops, LLM-driven code generation, RAG, validation and benchmarking of generated code
GPU / Performance	CUDA, kernel optimization, autotuning (KTT), Nsight Compute, NVRTC
Machine learning	PyTorch, computer vision, linear programming
Web	React

LANGUAGES

English **C1** German **B2** · DSD II certificate, B2/C1

SAILING

3rd place	29er EuroCup 2022
3×	Czech Republic champion
2022	Czech Junior Sailor of the Year

Former competitive skiff sailor: teamwork, boat handling, race preparation, and performance improvement.