Kseniya Parkhamchuk

k.parkhamchuk@gmail.com | https://www.kseniyaparkhamchuk.com/ | https://github.com/KsuParkhamchuk

Senior Fullstack Engineer with over 5 years of experience, actively transitioning into AI Engineering, focusing on LLMs. Leverages a strong software engineering foundation (best practices, system design) to develop AI-focused solutions. Valuing first-principle thinking and an integrated approach.

Skills

LLM / NLP: Python, PyTorch, FastAPI, vLLM Transformers, LoRA - QLoRA fine-tuning, custom datasets, prompt engineering, tokenization, Flash Attention, KV Cache, Paged Attention, RoPE, NCCL, FSDP, W&B.

Web Development: TypeScript / JavaScript, NextJS, React, NodeJS, MongoDB, AWS, Tailwind, SCSS, CSS Modules, REST API / GraphQL, Jest, React testing library, OAuth2, Custom Components Libraries, Redux Toolkit / TanStack Query, Storybook, pnpm, Vercel, Snyk, micro-frontends, Git

Projects

GPT-2 (https://github.com/KsuParkhamchuk/GPT-2)   PyTorch, KV Cache, RoPE, W&B
- Implemented GPT-2 architecture from scratch, including manual layer construction and algorithms implementation

vLLM-mm-chat (https://github.com/KsuParkhamchuk/vllm-mm-chat-comparison)   vLLM, FastAPI, WebSockets
- Architecting a chat system with vLLM, implementing parallel inference optimization and conducting comparative analysis against leading frameworks for optimal performance

Gemma3-sft (https://github.com/KsuParkhamchuk/gemma3-sft)   LoRA, Transformers, custom dataset
- Developed experimental fine-tuning pipeline, including custom data preprocessing for targeted model specialization

Experience

Senior Software Engineer, Kaseya   Jan 2024 – Dec 2024
- Led system design and initial implementation of monolith-to-microfrontends migration, enabling independent deployments and reducing cross-team dependencies across 4 teams while improving fault tolerance through granular release cycles twice per week.
- Delivered critical navigation system redesign in 48-hour sprint, resulting in successful product demo that improved feature usage from 25% to 47% of weekly active users and enabled a high-impact demo for high-profile stakeholders
- Spearheaded unification strategy for 2 distinct frontend portals, designing and implementing a shared component library with ~50 components that streamlined development across both platforms
- Optimized file system UI performance from ~3 sec to near-instant response time through custom state management and intelligent tree preloading, handling partial selection states across dynamically loaded nodes

Software Engineer, EPAM   Nov 2021 – Dec 2023
- Designed and delivered a comprehensive image processing solution, including tech stack selection and feature optimization for cropping, uploading, and resizing functionality
- Mentored 5 team members across several projects through targeted code reviews and technical knowledge transfers, strengthening frontend expertise across the team
- Led the end-to-end development of an MVP from scratch, including project architecture design, technology stack selection, and implementation planning, including cross-team collaboration

Software Engineer, Itransition   Oct 2020 – Oct 2021
- Developed diverse contract template system with multiple viewing modes (edit/preview), implementing custom HTML/CSS solutions to automate documentation generation across various business scenarios

Software Engineer, Softacom   Sep 2019 – Oct 2020
- Developed landing pages and content features while contributing to UX through design prototyping and SEO optimization, collaborating cross-functionally to deliver responsive, conversion-focused web experiences

Education

BSUIR – BS in Computer Science   June 2021