Sardaana Eginova

eginovasardaana@berkeley.edu | sardaana-eginova | daanaea.github.io | daanaea

Education

University of California, Berkeley **Graduation: May 2027**
B.A. Computer Science and B.A. Data Science (Applied Math emphasis) **GPA: 3.987/4.00**

Coursework: Deep Neural Networks, Data Structures, Discrete Math, Linear Algebra, Optimization, Probability Theory and Random Processes, Graph Theory, Operating Systems, Internet Architecture, Cybersecurity, Machine Structures, ASIC

Teaching: Head TA for DATA 100 (Principles of Data Science), Tutor for CS 70 (Discrete Math & Probability)

Awards: Dean's List (x4), WorldSkills Russia Big Data & ML Finalist, All-Russian Math Olympiad Regional Winner (2x)

Clubs: Machine Learning @ Berkeley, Cal Hacks, SAAS Data Science, Society of Women Engineers, Accel Scholars

Work Experience

Meta, Instagram Capacity team, Data Scientist Intern Jun 2026-Aug 2026

- Built a capacity management dashboard that identifies product teams approaching CPU quota, surfaces launches and experiments regressing capacity, and recommends levers to mitigate quota breaches
- Trained structural time series models to forecast web compute usage and flag at-risk surfaces, achieving 3x higher accuracy than linear projection baseline; validated with engineers, presented to stakeholders, integrated into tools

AWS Annapurna Labs, Software Engineer (Consultant) Sep 2025-May 2026

- Created API for Group-Relative Policy Optimization posttraining in Hugging Face on Trainium 2 chips via torch-xla
- Accelerated inference of video diffusion transformer, Wan 2.2, via PyTorch-native distributed inference: 4x longer video generation, 8.8x RoPE speedup with custom NKI kernel, 1.4x cross-attention speedup with prompt caching

inDrive, Marketplace Efficiency team, Data Analytics Intern Jul 2025-Aug 2025

Global #2 ride-hailing platform, 200M+ downloads across 45+ countries

- Tracked driver behavior from 2M+ daily app logs (accept/cancel rates, reaction times, ride distance preferences) to train CatBoost model for chained Smart Notifications dispatch
- Deployed switchback experiments in 10 cities with A/B testing and CUPED, achieving a 110% uplift in money-per-hour driver's metric in treatment group

Spotify, Machine Learning Engineer (Consultant) Feb 2024-May 2024

- Developed a customer support chatbot for Spotify's Freemium users for faster automated responses
- Analyzed 10,000+ text chats via TF-IDF topic modeling to synthesize more data, optimizing Q&A coverage by 120%
- Finetuned RAG (Mistral-7B, Gemini, Claude 3) with FAISS vector database by prompting and hyperparameter tuning

Yandex & North-Eastern Federal University, Researcher Jul 2021-Sep 2022

- Parsed 100,000+ word examples from news and forums, creating a training corpus capturing modern Yakut language
- Built a Russian-to-Yakut translator using contextual similarity and clustering techniques to enhance word-specific accuracy, contributing to low-resource language preservation for Siberian indigenous communities

Research

Berkeley Artificial Intelligence Research Lab (BAIR), Speech Group, Researcher Feb 2025-May 2025

- Ran ablation study for LLMs interpretability on causal tracing for transformer models (GPT-2XL, Llama 3, Phi-3)

Sky Computing Lab, Researcher Oct 2024-Jan 2025

- Created 100 Q&A BFCL benchmark for web-search and built end-to-end automated evaluation system for agentic LLMs testing web-search and function-calling using AST (exact match) and state-based metrics across 16 categories

UC Berkeley College of Computing, Data Science, and Society, Researcher Jan 2024-May 2024

- Built a 300K+ entries SPARQL knowledge graph from dental domain articles, reduced query response time by 60%

Skills

Languages: Python, C/C++, Java, JavaScript, PHP, Rust, SQL, R, SystemVerilog, RISC-V Assembly

Libraries: PyTorch, Tensorflow, OpenCV, XGBoost, Pandas, Scikit-Learn, NumPy, LangChain, React, Selenium

Tools: Git, Linux, Bash, CUDA, tmux, Docker, Kubernetes, Redash, gRPC/Protobuf, Claude Code/Cursor/Copilot