Sami Ansari

+1 (647) 289-6889 | SAnsari2805@Outlook.com | linkedin.com/in/samiansari99 | github.com/y0usaf

Experience

AI-Native Software Engineer   Oct 25 – Present
Cookunity, Toronto, Ontario, Canada
- Ship production features daily using AI-assisted development tools (Cursor, Windsurf, Claude Code), rapidly prototyping PM ideas from concept to working demo in hours
- Build internal tools, automations, and reusable LLM integration patterns to increase team velocity and democratize AI development across engineering and product teams

AI Engineer (Contract)   Oct 2024 – Dec 2024
Rootly, Toronto, Ontario, Canada
- Built proof-of-concept AI agent framework using RAG for automated root cause analysis with end-to-end data pipeline integrating Rootly, Slack, GitHub

Machine Learning Engineer Intern   Apr 2024 – Oct 2024
Cohere, Toronto, Ontario, Canada
- Enhanced LLM quantitative reasoning by **3%** through development of novel fully synthetic training dataset without human annotation, becoming the **first engineer to contribute a fully synthetic dataset** to Command-R-Plus foundation model
- Created specialized evaluations for length control and adapted FinanceBench for base model assessment
- Conducted model ablation studies analyzing impact of specific datasets on cross-category performance
- Collaborated cross-functionally with evaluations, multilingual, and safety teams to generate synthetic multilingual data for model training

Senior Data Quality Team Lead   Jun 2023 – Apr 2024
Cohere, Toronto, Ontario, Canada
- Designed and deployed scalable data labeling platform with React/TypeScript frontend and Firebase/Firestore backend, processing 100,000+ tasks used by 200+ labelers, **saving $3M in operational costs**
- Built backend data infrastructure and developed Python scripts with internal AppsScript applications analyzing tasks to extract insights for ML model improvements
- Led SFT/RLHF labeling projects and first cohort of STEM and code-focused labelers, working with external data vendors to consolidate data quality standards for post-training

Data Quality Specialist   May 2023 – Jun 2023
Cohere, Toronto, Ontario, Canada
- Delivered high-impact annotations including preference ranking, quality scoring, and SFT conversation generation across all domains

Technical Skills

Languages: Python, TypeScript/JavaScript, SQL, C/C++, Bash, Nix
ML & AI: LLM Fine-tuning, RAG, RLHF, SFT, Synthetic Data Generation, Model Evaluation, PyTorch, JAX
Frameworks & Libraries: FastAPI, React, Pandas, Polars, Tailwind
Infrastructure: AWS (S3, Lambda, Bedrock), Docker, Linux, NixOS, Slurm, Git, Redis, DuckDB, Firebase, Firestore

Speaking

Guest Lecturer on AI and Synthetic Data Generation for MMA Program   2024
Queen's Smith School of Business, Kingston, Ontario | 50+ graduate students

Projects

NixOS System
- Comprehensive system configuration for Desktop, Laptop, and Server environments with Nix-Darwin support for macOS
- Integrated services including Syncthing for file synchronization and Forgejo for self-hosted Git