Sami Ansari

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

Experience

AI-Native Software Engineer
Cookunity, Toronto, Ontario, Canada
Oct 2025 – Present
- Forward-deployed across engineering, product, customer service, product tagging, and performance review teams using Cursor, Windsurf, and Claude Code
- Created internal plugins repository and reusable LLM integration patterns, democratizing AI development across the organization

AI Engineer (Contract)
Rootly, Toronto, Ontario, Canada
Oct 2024 – Dec 2024
- 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
Cohere, Toronto, Ontario, Canada
Apr 2024 – Oct 2024
- 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
Cohere, Toronto, Ontario, Canada
Jun 2023 – Apr 2024
- 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
Cohere, Toronto, Ontario, Canada
May 2023 – Jun 2023
- 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
Queen's Smith School of Business, Kingston, Ontario | **50+ graduate students**
2024

Projects & Contributions

- **NixOS System** – Built multi-device config (Desktop, Laptop, Server, macOS) with self-hosted services
- **tweakcc** – Contributed bunx support to Claude Code customization tool
- **claude-code-nix** – Contributed hourly builds and Cachix binary cache integration