Aritra Ghosh

(858) · 214 · 7058 ◇ aritra.io ◇ github/arighosh05 ◇ in/ghosh-aritra ◇ aritraghosh534@gmail.com

Education

University of California, San Diego — September 2022 - June 2025
B.S. in Computer Science — *San Diego, CA*

- **Relevant Coursework**: Statistical NLP, Recommender Systems, Web Mining, Software Construction

Work Experience

Abu Dhabi National Oil Company — Summer 2024
Machine Learning Engineer Intern — *Abu Dhabi, UAE*

- Developed robust machine learning pipeline for oil spill detection utilizing hyperspectral satellite imagery
- Diversified data via GAN-based augmentation & RGB-to-HSI conversion, boosting performance by 15%
- Automated oil spill mapping, reducing manual analysis time by 70% through visual segmentation models
- Engineered forecasting system using environmental data to forecast the movement & impact of oil spills
- Optimized model deployment via cloud-based processing, model quantization, and parallelized inference
- Prototyped live monitoring system with automated alerts, amortizing detection-to-response time by 50%

Projects

Coding Agent | [*Code*] [*Blog*]

- Developed agentic code review framework with augmented LLM calls & human-in-the-loop mechanisms
- Designed parallelized LLM calls for static code analysis & planner-executor workflow for code generation
- Integrated context-aware tool-selection & tool-calling with customized functions for static code analysis
- Implemented RAG pipeline via semantic search on vector database for contextual documentation retrieval

Web Crawler | [*Code*] [*Blog*]

- Devised differentially private web crawler to synthesize analytically valuable data from extracted content
- Engineered text transformation system utilizing semantic preservation algorithms & statistical resampling
- Added configurable privacy budget system allowing users to balance data utility vs. privacy guarantees
- Created parallel processing system to enable concurrent processing of multiple web sources transformation

Model Probe | [*Code*] [*Blog*]

- Engineered monitoring system that intercepts & analyzes intermediate CoT reasoning steps from LLMs
- Built sentiment analysis pipeline to detect value misalignments between AI systems & human stakeholders
- Incorporated rule-based checks for known policy violations & unsupervised learning for anomaly detection
- Designed real-time alerting system to notify stakeholders when model deviates from intended value system

Prompt Silo | [*Code*] [*Blog*]

- Built secure offline-first private prompt management system & deployed as Obsidian community plugin
- Implemented timestamp logging, PII redaction, & AES-256 encryption to store prompts with metadata
- Developed dual-key system with separate access tiers, enforcing fine-grained controls to facilitate visibility
- Integrated automated local backup & archival process, preventing data loss & enforcing data redundancy

Technical Skills

Certifications	AWS AI Practitioner, AWS Machine Learning Engineer, AWS Data Engineer
Languages	Python, Java, C++/C, Julia, Scala, Go, R, MATLAB, SQL, Perl, JavaScript
Software	PyTorch, TensorFlow, LangChain, LlamaIndex, Docker, Kubernetes, Jenkins