RISHIK SARKAR

New York, NY | +1-732-783-8669 | rishiksarkar02@gmail.com | github.com/RishikSarkar | rishiksarkar.com

EDUCATION

Cornell University	Aug 2024 – May 2025
Master of Engineering in Computer Science	GPA: 3.9/4.0
Rutgers University-New Brunswick	Sep 2020 – May 2024
Bachelor of Science in Computer Science (Honors), Cognitive Science	GPA: 3.9/4.0
EXPERIENCE	

Machine Learning Engineer

Instalily AI (Google AI Accelerator)

- Built a RAG-based agent and MCP server for United Rentals, leveraging Azure AI Search for semantic retrieval and a custom LangChain-style framework integrated with UR APIs.
- Optimized embedding pipelines and vector index queries, reducing average retrieval latency by 40%.
- Refined the frontend chat agent and added **PostgreSQL** logging and metrics for monitoring and optimization.

Independent Researcher

Cornell XR Collaboratory

- Spearheading ML-driven AR/VR research for Quest platforms under Prof. Harald Haraldsson, developing a Unity package with novel 3D interaction techniques using AR Foundation, DOTWeen, and the XR Interaction Toolkit.
- Leveraging Microsoft.Extensions.AI and Ollama LLMs in a .NET microservice architecture to automate object selection/manipulation and power an automated mind-map system, accelerating XR workflows.

ML Full-Stack Developer Intern

Provenir (Fintech)

- Developed an automated credit-risk decisioning system by integrating Decision Trees, Random Forests, XGBoost, and RNNs via scikit-learn and TensorFlow, achieving 95% accuracy.
- Enhanced model interpretability with SHAP and LIME, wrote 100+ MockMvc unit tests, and containerized deployments on Minikube, boosting reliability by 20% and reducing decision latency by 98%.
- Optimized API endpoints for artifact generation and log retrieval to enable real-time monitoring, driving a 135% increase in conversions.

ML Research Intern

Abraira Lab

- New Brunswick, NJ Preprocessed and curated 10,000+ behavioral samples with Motion Sequencing (MoSeq2) in Python, enabling robust unsupervised modeling in a neuroethology study.
- Corrected anomalous keypoint detections, improving data quality by 60% and boosting model accuracy.

PROJECTS

MiniTorch | Python, PyTorch, CUDA, Numba

- Reimplemented the Torch API from scratch, including autodifferentiation, broadcasting, and gradient ops for robust backpropagation.
- Built a custom tensor library enabling multi-dimensional operations, parallelized with CUDA and Numba for high-performance computing.

Protoclear | Next.js, FastAPI, TF-IDF, NER, Chroma

- Developed an IRB compliance toolkit using **TF-IDF** for keyword extraction and rule-based **NER** to flag research-specific terms.
- Integrated LlamaIndex with a Chroma vector store to implement a RAG pipeline for regulatory guidance.

TECHNICAL SKILLS

Languages: Python, Java, C#, C++, JavaScript, TypeScript, C, SQL

Frameworks/Libraries: PyTorch, TensorFlow, scikit-learn, Keras, LangChain, LlamaIndex, RAG, OpenCV, Pandas, CUDA, Numba, Flask, FastAPI, Next.js, Tailwind CSS, Beautiful Soup, Tkinter, JUnit, MockMvc, MongoDB, MySQL, PostgreSQL, SQLite, Unity (UPM, XR Interaction Toolkit), .NET

Tools/DevOps: Docker, Kubernetes, Minikube, Amazon AWS, Azure Search, Google Cloud Platform, Jenkins, Git, Jupyter, CI/CD pipelines, Jira, MCP

May 2022 – Jun 2023

Aug 2024 – Dec 2024

Aug 2024 – Dec 2024

New York, NY

Mar 2025 - Present

Jan 2025 – May 2025

New York, NY

Jun 2023 - Dec 2023 Parsippany, NJ