

# RISHIK SARKAR

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

## EDUCATION

### Cornell University

Master of Engineering in Computer Science

Aug 2024 – May 2025

GPA: 3.9/4.0

### Rutgers University-New Brunswick

Bachelor of Science in Computer Science (Honors), Cognitive Science

Sep 2020 – May 2024

GPA: 3.9/4.0

## EXPERIENCE

### Machine Learning Engineer

Instalily AI (Google Startup Accelerator)

Mar 2025 – Present

New York, NY

- Built an agent for United Rentals using **RAG**, semantic search, **Azure AI Search**, **MCP**, and a custom LangChain-like framework integrated with UR APIs.
- Optimized embedding generation and vector database queries, boosting retrieval speed by over 40%.
- Refined the frontend chat agent and added **PostgreSQL** logging and metrics for monitoring and optimization.

### Independent Researcher

Cornell XR Collaboratory

Jan 2025 – May 2025

New York, NY

- Spearheading ML-driven AR/VR research for Quest platforms under Prof. Harald Haraldsson, developing a Unity package with novel 3D interaction techniques using **Unity**, **UPM**, and the **XR Interaction Toolkit**.
- Leveraging **Microsoft.Extensions.AI** and **Ollama** LLMs within 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)

Jun 2023 – Dec 2023

Parsippany, NJ

- Developed an automated credit risk decisioning system by integrating multiple ML models (Decision Trees, Random Forests, XGBoost, RNNs) with **scikit-learn** and **TensorFlow**, achieving 95% accuracy.
- Enhanced model transparency with **SHAP/LIME**, wrote 100+ unit tests (**MockMvc**), and streamlined **Minikube** deployments, boosting reliability by 20% and reducing time-to-decision by up to 98%.
- Refined API endpoints for artifact generation and log retrieval to enable real-time monitoring, contributing to a 135% increase in conversions.

### ML Research Intern

Abraira Lab

May 2022 – Jun 2023

New Brunswick, NJ

- Preprocessed and curated over 10,000 training samples using **Motion Sequencing (MoSeq2)** (Python) for an unsupervised behavioral model in a neuroethology study.
- Rectified anomalous key point detections, improving data quality by 60% and strengthening model performance.

## PROJECTS

### MiniTorch | Python, PyTorch, CUDA, Numba

Aug 2024 – Dec 2024

- 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

Aug 2024 – Dec 2024

- Developed a compliance toolkit for IRB regulations, leveraging **TF-IDF** for keyword extraction and rule-based **NER** to flag research-specific terms.
- Integrated **LlamaIndex** with a **Chroma** vector database, creating a **RAG** workflow 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