

RISHIK SARKAR

+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

Rutgers University-New Brunswick

Sep 2020 – May 2024

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

GPA: 3.9/4.0

- Honors: Summa Cum Laude, SAS Honors Program, Phi Beta Kappa, Dean's List

EXPERIENCE

ML Full-Stack Developer Intern

Jun 2023 – Dec 2023

Provenir (Fintech)

Remote

- Built an automated credit risk decisioning solution by integrating Decision Trees, Random Forests, XGBoost, and RNNs into FLAML using **scikit-learn** and **TensorFlow**, achieving a 95% prediction accuracy in customer credit risk assessments through hyperparameter tuning and monotonic constraints
- Collaborated with a team of 5 engineers to enhance AI explainability by incorporating SHAP and LIME visualizations, enabling stakeholders to understand the rationale behind risk scores and make more informed, data-driven decisions in real-time
- Implemented over 100 unit tests with **MockMvc**, increasing software reliability by 20%, and streamlined model deployment on **Minikube**, contributing to scalable testing practices adopted in subsequent releases
- Refined API endpoints for artifact generation and log retrieval, which enabled seamless monitoring of model performance and continuous learning from data to optimize decision-making processes

ML Research Intern

May 2022 – Jun 2023

Abraira Lab

Hybrid

- Employed **Motion Sequencing** (MoSeq2) to preprocess and create a dataset of over 10,000 high-quality training samples for an unsupervised learning model analyzing behavioral syllables in mice with spinal cord injuries for a computational neuroethology study
- Analyzed behavioral syllables identified by the model and rectified anomalous keypoint results, leading to a 60% improvement in data quality

PROJECTS

MiniTorch | Python, PyTorch, CUDA, Numba

Aug 2024 – Present

- Developing a re-implementation of the **Torch API**, as part of Cornell's Machine Learning Engineering course
- Implemented autodifferentiation from scratch, enabling backpropagation and gradient computations essential for training neural networks
- Constructed a custom tensor library, creating data structures and operations to handle multi-dimensional arrays, crucial for deep learning computations
- Optimized performance through parallel computing, utilizing **CUDA** for GPU acceleration and **Numba** for Just-In-Time (JIT) compilation, enhancing computational efficiency
- Built foundational neural network components, including layers, activation functions, and loss functions, ensuring compatibility with **PyTorch** models and facilitating seamless integration

Invasion of the Bot-Grabbers | Python, Jupyter, Pandas, PyTorch, Matplotlib

Sep 2023 – Dec 2023

- Implemented search algorithms including **A*** and **D* Lite** to develop an automatic crew rescue maze simulation for a graduate-level course
- Enhanced the bot's decision-making by integrating **Bayesian networks**; utilized sensors to determine optimal paths to crew members probabilistically while avoiding moving obstacle aliens
- Trained two **logistic regression** models to predict the bot's moves and win probabilities, incorporating features engineered from the probability matrices calculated earlier, enhancing efficiency
- Advanced the bot's performance by implementing an ACTOR-CRITIC RL framework with **PyTorch**

TECHNICAL SKILLS

Languages: Python, Java, JavaScript, C++, C, C#, Kotlin, Dart, Rust, SQL, MATLAB, Scheme

Frameworks and Libraries: PyTorch, TensorFlow, scikit-learn, Keras, OpenCV, Pandas, Flask, Spring, Next.js, Tailwind CSS, Beautiful Soup, Tkinter, JDBC, CUDA, Numba, JUnit, MockMvc, MongoDB, MySQL, SQLite

DevOps and Tools: Docker, Kubernetes, Minikube, Jenkins, Amazon AWS, Git, Jira, Jupyter