

RISHIK SARKAR

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

EDUCATION

Cornell University

Master of Engineering in Computer Science

New York, NY

Aug 2024 - (May 2025)

- *Relevant Coursework:* Applied Machine Learning, Machine Learning Engineering, Virtual and Augmented Reality, HCI and Design, Introduction to Computer Vision, Natural Language Processing, 3D Interaction Design, Machine Learning for Health, Productizing Machine Learning

Rutgers University-New Brunswick

Bachelor of Science in Computer Science, Cognitive Science

New Brunswick, NJ

Sep 2020 - (May 2024)

- GPA: 3.9/4.0
- Honors: Summa Cum Laude, SAS Honors Program, Phi Beta Kappa, Dean's List
- *Relevant Coursework:* Data Structures, Design and Analysis of Computer Algorithms, Principles of Programming Languages, Numerical Analysis and Computing, Principles of Data Management, Formal Languages and Automata, Introduction to Artificial Intelligence (Graduate)

TECHNICAL SKILLS

- **Languages:** Python, Java, JavaScript, C++, C, C#, Kotlin, Dart, Rust, SQL, MATLAB, Scheme, Swift
- **Web Technologies:** HTML, CSS, Flask, Next.js, Tailwind CSS, Spring
- **Frameworks and Libraries:** PyTorch, TensorFlow, scikit-learn, Keras, OpenCV, Beautiful Soup, JavaFX, Pandas, Tkinter, NumPy, Python NLTK, SpaCy, Hadoop, Unity
- **Databases:** MySQL, MongoDB, SQLite, JDBC
- **DevOps and Tools:** Docker, Kubernetes, Jenkins, AWS, Git, Jira, Jupyter
- **Fields and Concepts:** Computer Vision, Natural Language Processing (NLP), Data Science, Data Structures & Algorithms

INDUSTRY EXPERIENCE

Software Engineer (Part-time)

AllAboutID

Oct 2024 - Present

Remote

- Built a secure and streamlined asset management system for an interior design startup, leveraging **MongoDB Atlas** and SVG handling for seamless furniture integration within the API
- Developed the frontend using **Next.js** and **Tailwind CSS**, including asset display features, enhancing the user interface and user experience

ML Full-Stack Developer Intern

Provenir

Jun 2023 - Dec 2023

Parsippany, NJ

- 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

SAT/ACT Tutor

The Princeton Review

Feb 2023 - Jun 2023

Princeton, NJ

- Utilized innovative teaching technologies and instructional design principles to motivate students and promote academic excellence
- Frequently stayed after hours to address individual student queries, ensuring thorough understanding and boosting confidence in test preparation

RESEARCH EXPERIENCE

Research Assistant

Sep 2023 – Aug 2024

Princeton University (CCNP)

Hybrid

- Created **Python** scripts to transform 800+ Excel, CSV, and JSON files from five clinical studies into a consolidated **SQLite** database, seamlessly integrating automated schema generation with key constraints. The data encompasses clinical assessment results for mental health conditions like depression and BPD, using scales and structured interviews (e.g., HDRS, BPRS, PANSS, SCID)
- Designed a streamlined **Tkinter**-based GUI to simplify database interactions for researchers without technical knowledge, incorporating advanced functionality for executing custom SQL queries through **Pandas**

ML Research Intern

May 2022 – Jun 2023

Abraira Lab

New Brunswick, NJ

- Employed Motion Sequencing to preprocess and create a dataset of over 10,000 high-quality training samples for an unsupervised ML model in a Computational Neuroethology observation study
- Analyzed behavioral syllables identified by the model and rectified anomalous keypoint results, leading to a significant 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

- Developed a grid maze simulation for a graduate-level course, Introduction to Artificial Intelligence (16:198:520), with a bot navigating to save crew members while avoiding aliens
- Implemented search algorithms including BFS, A*, and D* Lite to facilitate the bot's path-finding through the maze during the first iteration of the project
- Conducted extensive tests by varying the number of aliens and grid sizes, visualizing and analyzing the bot's performance metrics regarding its survivability and the number of crew members it saved
- Enhanced the bot's decision-making in the second iteration by integrating Bayesian networks; utilized sensors to determine optimal paths to crew members probabilistically
- Trained two logistic regression models to predict the bot's moves and win probabilities in the third iteration, incorporating features engineered from the probability matrices calculated earlier, enhancing efficiency
- Advanced the bot's performance further by implementing reinforcement learning with PyTorch, using an ACTOR-CRITIC framework. The ACTOR network predicted the bot's next move and the CRITIC network predicted the alternate move with the highest probability of success, enabling iterative improvement

Tch.ai | Next.js, Tailwind, Flask, Keras, OpenCV, Pandas, MySQL

Apr 2023 – Jul 2023

- Innovated a full-stack web application to deploy a Keras image classifier and tokenizer that recommend songs based on mood predictions from facial expressions or textual data
- Trained the classification model on the FER-2013 dataset and utilized OpenCV and a Haar Cascade classifier to preprocess datasets: achieving a training accuracy of around 96% and a validation accuracy of over 70%
- Designed a Next.js and Tailwind frontend that supports three genre selection methods—image, text, and manual choice—thus elevating user engagement with seamless data uploads and song recommendations
- Crafted a Flask REST API backend for image data preprocessing, providing personalized playlists from a CSV of 114,000+ songs, fine-tuned using genre, mode, valence, and other features
- Integrated a remote MySQL database, allowing users to create accounts and manage liked songs

LEADERSHIP AND SERVICE

Product Studio Collaborator

Sep 2024 – Present

Weill Cornell Medicine, Cornell University

New York, NY

- Collaborating with Weill Cornell Medicine (CWID: ris4016) as part of Cornell's Product Studio course to develop solutions for health tech entrepreneurs and researchers to streamline compliance with HIPAA and IRB guidelines

Chief Technology Officer

Sep 2023 – Aug 2024

Health Model United Nations

New Brunswick, NJ

- Orchestrated the design, development, and maintenance of the conference website utilizing Next.js and Tailwind CSS, ensuring real-time updates and a user-friendly interface
- Collaborated closely with other executive board members, playing a pivotal role in establishing communication infrastructures and social media platforms, as well as spearheading fundraising ventures
- Participated actively in the recruitment process, assisting in interviewing and evaluating potential board and staff candidates to uphold the standards of the conference

Co-Founder

Mar 2022 – May 2023

The Verbose Project

New Brunswick, NJ

- Developed a non-profit tutoring platform catering to incoming and first-year college undergraduates, focusing on fostering academic growth and empowerment
- Produced and released a series of "Introduction to Python" tutorial videos, which garnered positive feedback and enriched the learning experiences of students
- Provided supplemental tutoring assistance to students at Rutgers learning centers, augmenting the resources available to them and ensuring a solid foundational understanding

EXTRACURRICULAR AFFILIATIONS

- *The Phi Beta Kappa Society* - Member, 2023
- *Undergraduate Student Alliance of Computer Scientists* - Member, 2021 - 2023
- *Cognitive Science Club* - Member, 2021 - 2023

CERTIFICATIONS

- Human Subjects Research (IRB) Training, CITI Program - Completed Sep 2024 (Record ID: 64849869)
- *Deep Learning Specialization*, DeepLearning.AI (Coursera) - Completed Oct 2022
- *Machine Learning Specialization*, Stanford University (Coursera) - Completed Jun 2022
- *AI For Everyone*, DeepLearning.AI (Coursera) - Completed Mar 2022

PUBLICATIONS / ARTICLES

- Eisdorfer, J.T., Thackray, J., Theis, T., Vivinetto, A., Ricci, M.T., Lin, S., Oputa, O., Martinez, A.M., Nacht, H.D., Tschang, M., Mahmood, M., Tucker, A., Bohic, M., Pusuloori, S., Zmoyro, L., Abaira Lab Computational Group, Popovich, P., Ferguson, A.R., McTigue, D., Tysseling, V.M., Dulin, J., Hollis II, E., Datta, S.R., Abaira, V.E. (2023). *The Behavior Biomarker Scale (BBS): A machine-vision approach for automated locomotor recovery evaluation at millisecond timescales*. bioRxiv. <https://doi.org/10.1101/2023.10.31.564826>
- Sarkar, R. (2023, August 2). *Student insights: The Dark Side of Chatbot Therapy*. Critical AI. [Online]. Available: <https://criticalai.org/2023/07/31/student-insights-the-dark-side-of-chatbot-therapy/>

LANGUAGES

- English (Native)
- Bengali (Native)
- Hindi (Advanced)
- Japanese (Intermediate)
- German (Beginner)