Roshan Kenia

roshan.kenia@stonybrook.edu | linkedin.com/in/roshan-kenia/ | github.com/roshankenia

Education

Stony Brook University - Stony Brook, New York   **Expected Graduation: May 2023**
Bachelor of Science in Computer Science, Applied Math and Statistics   Dean's List
GPA: 4.0/4.0   University Scholars Honors Program

Relevant Courses

Machine Learning, Natural Language Processing, Data Science, Data Mining, Analysis of Algorithms, Data Structures and Algorithms, Object-Oriented Programming (Java), Linear Algebra, Graph Theory, Software Engineering, Computer Networks

Technical Skills

- Languages: Java, Python, JavaScript, C, HTML, CSS, MySQL
- Software Engineering Technologies: MongoDB, Express, React, Node.js (MERN)
- Machine Learning Technologies: Tensorflow/Keras, Pytorch, Numpy, Pandas, Matplotlib, Seaborn, Sklearn

Work Experience

Stony Brook University, Stony Brook, NY — *Machine Learning Researcher*   **January 2022 - Present**
- Studying with Dr. Zhaozheng Yin the problem of learning under the presence of noisy labels within a dataset.
- Mitigating ways in which a robust model can be built so as to be tolerable to overfitting caused by noisy data.
- Utilizing semi-supervised learning by treating noisy samples as unlabeled in order to compete with the state-of-the-art.
- Continued my work full time as a URECA Machine Learning Program Scholar in the summer of 2022.

SupplyHouse.com, Melville, NY — *IT Intern*   **June 2021 - July 2021**
- Applied Scrum methodologies whilst working on projects for the e-commerce company's backend admin system used by other departments such as Forecasting, Customer Service, and Buying.
- Developed new querying methods in Java with the use of MySQL that decreased total query time by 83%.

Koo Laboratory, Cold Spring Harbor, NY — *Machine Learning Researcher*   **September 2020 - May 2021**
- Researched the functional impact of genomic mutations through a computational lens using machine learning solutions.
- Conducted a two semester long project involved with distilling knowledge from a teacher network to a student network as a regularization technique.
- Showed during a laboratory meeting student-teacher networks could produce more interpretable results by up to 20%.

Projects

Uncertainty Metrics for Learning With Noisy Labels / Pytorch, Sklearn, Seaborn   **June 2022 - August 2022**
- Investigated ways uncertainty metrics may be used to separate noisy samples from clean ones in a corrupted dataset.
- Found metrics such as entropy and peak ratio could be employed while training to identify noise with an 80% accuracy.
- Recognized using dimensionality reduction algorithms such as tSNE and PCA greatly assisted in better visualizations of noisy sample vectors in a 2D or 3D space.

JART: Joining Artists in Real Time / MongoDB, Express, React, Node.js   **January 2022 - May 2022**
- Designed a comic and story building game site in which users could play together, make and interact with posts, chat with other users, and discover new communities for topics they were interested in.
- Created and queried a MongoDB server that stored users posts, friends, likes, dislikes, comments, and communities.
- Used socket.io to allow users to see what others drew or wrote in real time while playing the game.

Knowledge Distillation in Genomic Deep Learning / TensorFlow, Keras, Pandas   **December 2020 - May 2021**
- Researched the process of knowledge distillation which involved training a student network on a complex teacher network using DNA sequences as data.
- Showed that student networks can be used as more interpretable and general models compared to their teacher networks through the use of saliency maps, integrated gradients, and in silico mutagenesis.