Samraj Moorjani

919-939-4031 | samrajmoorjani@gmail.com | smoorjani.github.io

EDUCATION

University of Illinois at Urbana-Champaign

GPA 4.00/4.00

Master of Science in Computer Science

Thesis Advisor: Hari Sundaram

Urbana, IL | August 2022 - May 2023

Bachelor of Science in Computer Science

Urbana, IL | August 2019 - May 2022

William G. Enloe High School

GPA 4.7/4.0 (Weighted)

Took Calculus III and Differential Equations at NC State University.

Raleigh, NC | August 2015 - June 2019

SKILLS

Languages: Python, C++, CUDA, Java, LaTeX, TypeScript, HTML, CSS, MATLAB, Haskell

ML Libraries: Huggingface, PyTorch, W&B, TensorFlow, OpenCV, Keras

Databases: SQL, Spark, Presto, Hive, BigQuery, Vertica, PostgreSQL, MongoDB, Neo4j

DevOps: Airflow, Jenkins, Kubernetes, AWS EC2, GCP, Git, Docker, Shell,

HONORS

Siebel Scholar - Class of 2023

C3SR URAI Scholar

First Place Caterpillar Anomaly Detection in IoT Challenge

UIUC Dean's List

2x American Invitational Mathematics Exam Qualifier

COURSEWORK

Machine Learning

Adv. Natural Language Processing

Deep Learning

Knowledge-Driven Natural Language Generation

Statistical Reinforcement Learning

Computational Advertising

Database Systems

EXTRACURRICULAR

EMNLP, ACL Reviewer

SIGKDD, CIKM Subreviewer

HackIllinois 2021 Outreach Team

EXPERIENCE

Applovin, Palo Alto, CA — Machine Learning Engineering Intern

June 2023 - August 2023

- Worked on the Ad-Tech Research Science team to automate model & data pipeline health validation. Wrote pipelines to check for data/model drift and ensure online-offline consistency.
- Incorporated creatives embeddings from ResNet into install rate/ranking models and investigated other modalities (e.g., video and text). Helped build infrastructure to support creatives ranking.

Meta, Menlo Park, CA — Software Engineering Intern

May 2022 - August 2022

- Worked on Ads Responsibility and Privacy to improve multimodal advertisement classification using novel hierarchical multi-label classification methods.
- Improved efficiency and scalability of the model training workflow for other ML Engineers.

Facebook, Menlo Park, CA — Software Engineering Intern

May 2021 - August 2021

- Worked on AI Commerce Multimodal to improve the cross-lingual abilities of multilingual multimodal categorization models. Improved the accuracy of low-performance languages by 13% on average.
- Researched and implemented strategies such as iterative weak label self-learning, confidence thresholding, and text augmentation. Improved overall model performance by 3.6%

Teaching Assistant, University of Illinois at Urbana-Champaign

- **Computational Advertising** (Spring 2022, Spring 2023) - Responsible for creating worksheets, homeworks, and PrarieLearn problems, along with grading, and logistics.
- **Applied Machine Learning** (Fall 2022) - Held office hours and answered questions on Campuswire.
- **Algorithms** (January 2021 - December 2021) - Held office hours and assisted in grading homework.
- **Software Design Studio** (January 2020 - December 2020) - Moderated code reviews and provided feedback to students. Emphasized unit testing, proper OOP design, clean coding principles, etc.

Capital One, Champaign, IL — Software Engineering Intern

June 2020 - August 2020

- Deployed a distributable named entity recognition (NER) processor with a BERT-BiLSTM-CRF model.
- Greatly outperformed production-grade NER tools like spaCy and Flair.
- Introduced compatibility with other NLP architectures and provided parallel computing with Dask.

RESEARCH

Crowd Dynamics Lab (May 2021 - Present) - *Graduate Researcher*

- Advisors: Hari Sundaram, Ewa Maslowski, Adit Krishnan
- Employing natural language generation to present scientific information to a layperson audience
- Utilizing an adversarial, reward-based framework to distill into GPT-2 knowledge about traits that are difficult to quantify but easy to observe from examples (e.g. persuasiveness and memorability).

IBM C3SR (August 2020 - Present) - *Undergraduate Researcher in AI*

- Advisors: Vikram Mailthody Sharma, Mert Hidayetoglu, Wen-mei Hwu, Jinjun Xiong
- Facilitating faster sparse matrix operations with GPUs. We have integrated cuBLAS and cuSPARSE baselines into a custom kernel pipeline for PyTorch and are integrating our group's kernel.
- Pruned large language models like NVIDIA's Megatron-LM BERT with various structured/unstructured pruning techniques and retrained with weight rewinding.

PROJECTS

Chrome Dino Genetic Algorithm (Winter 2019)

- Created an algorithm from scratch in C++ to develop random genetic mutations based on automatically-defined selection criteria to ensure desirable populations are created.