

Reetahan Mukhopadhyay

194 Scholes Street, Brooklyn, NY 11206 | (781)-835-5080 | reetahan.m@columbia.edu
website: reetahan.com | github: @reetahan | linkedin: @reetahan-mukhopadhyay

EDUCATION

Columbia University

New York, NY

M.S. in Computer Science, GPA: 3.97

Dec 2022

- **Key Courses:** Unsupervised Learning, Causal Inference, Machine Learning & Climate, Networks & Crowds
- **Thesis:** The Effects of Homophily and Influence on Voting Games, advised by Prof. Augustin Chaintreau
- **Teaching Assistantships:** Lead TA for Analysis of Networks and Crowds | TA for Intro to Databases
- **Research Assistantship:** Data Science Institute/Lamont-Doherty Earth Observatory: Robustness of ML Pipeline to Project Aerosol Concentrations Against Missing Data & Customized Data Regridding Frameworks
- **Projects:** Causal Attribution: Climate Change & Hurricanes | Probabilistic Agglomerative Clustering for Redistricting

University of Illinois at Urbana Champaign

Urbana, IL

B.S. in Computer Science, GPA: 3.77

May 2021

- **Key Courses:** Machine Learning, Artificial Intelligence, Data Mining, Database Systems, Simulation, Operating Systems
- **Honors:** Professional Development Scholarship Winner (Columbia) • James Scholar (UIUC) • Dean's List (UIUC)

PROFESSIONAL EXPERIENCE

Virtualitics

Pasadena, CA

Data Scientist [Full-Time]

Apr 2022 - present

- Advancing solution flows for US DoD using automated theorem prover for constraint optimization.

Machine Learning Engineering Intern

- Developed automated ontology generation for US DoD using topic modeling, entity extraction & other NLP techniques upon PDF documents, and deployed network graph visualizations in software. Derived courses of action that aided in potential savings of over \$100 million annually using further data analysis on graph.
- Evolved AI at Scale initiative for ML flow generation (supporting a wide array of data sources) using distributed frameworks and online algorithms for anomaly detection tasks; built testing suite for the initiative.

Mavenir

Reading, MA

Artificial Intelligence Intern

May 2020 – Aug 2020

- Developed internal tools to detect errors in calculated key performance indicators (KPIs) and possible discrepancies in calculations performed between different divisions within the company.
- Utilized Elasticsearch-Logstash-Kibana stack to migrate data and KPIs, and create accessible visualizations of the log data. Created various additional scripts for task automation (all in Python).

Cadence Design Systems

Chelmsford, MA

Software Engineering Intern

May 2019 - Aug 2019

- Identified areas in generated code, used to execute Verilog / System Verilog code in a parallel logic simulator, that has poor performance with the instruction cache using performance analysis and profiling.
- Implemented a faster and memory efficient string hashing and management scheme & produced a GDB routine to aid in debugging source code (all in C/C++).

RESEARCH EXPERIENCE

Social Mobile Lab, Columbia

Feb 2022 - present

- Investigated the effect of multi-hop job referral schemes for minority groups in professional social network systems compared to traditional single-hop methods, *Presented at EAAMO 2022, under review for journal publication*
- Thesis research on the effect of homophily and influence threshold parameters on the outcome of an election in a novel interacting particle system model, through simulation and mathematical analysis.

Riemer Group, University of Illinois

June 2020 - Dec 2021

- Integrated use of newly observed data to evaluate and improve supervised learning models using simulated training data from PartMC to predict aerosol mixing state. *Presented at AAAR 2021, under review for journal publication*
- Conducted sensitivity analysis of the model, investigating effects of feature engineering and training set manipulation - including training data generated from my implementation of temperature variation for PartMC itself.

TECHNICAL SKILLS

Languages: Python • C • C++ • Java • SQL • Cypher • JavaScript • HTML / CSS • Ocaml • Clojure • Go • Latex

Tools/Libs: PyTorch • TensorFlow • AutoML • Spark • Scikit • Dask • Pandas • MongoDB • Databricks • Elastic Stack • Z3

Environments: AWS • Git • Docker • Jenkins • Firebase • Confluence • Android • Linux