Reetahan Mukhopadhyay

301 St Nicholas Avenue, New York, NY 10027 | (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 *(Fall 22)* | TA for Intro to Databases *(Fall 21)*
- ***Research Assistantship**:* Data Science Institute/Lamont-Doherty Earth Observatory *(Oct 21 - July 22)*: 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*
Machine Learning Engineering Intern *Apr 2022 - Aug 2022 | Jan 2023 - present*

- Developed automated ontology generation for DoD customer using topic modeling, entity extraction & other NLP techniques upon scattered PDF documents, and deployed network graph visualizations in software.
- Utilized this with further data analysis and graph algorithms to derive key insights and courses of action that aided in potential savings of over $100 million annually.
- Evolving AI at Scale initiative using distributed frameworks and online algorithms for anomaly detection and other ML routines, supporting a wide array of data sources, for data analysis flow generation; plus building initiative's testing suite.

Mavenir *Reading, MA (Remote)*
Artificial Intelligence Intern *May 2020 – Aug 2020*

- Developed internal tools (in Python) 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 (ELK) stack to migrate data and KPIs, and create accessible visualizations of the log data. Created various additional scripts for task automation.

Cadence Design Systems *Chelmsford, MA*
Software Engineering Intern *May 2019 - Aug 2019*

- Worked on the identification of areas in generated code used to execute Verilog/System Verilog code in a parallel logic simulator, that have poor performance with the instruction cache, using performance analysis and profiling
- Implemented a faster and reduced memory usage 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, through mathematical analysis & simulation on social network data. *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

• Python • SQL • Git • C / C++ • Linux • AutoML • Spark / PySpark • PyTorch • TensorFlow • Dask • NumPy / Pandas
• Java • Latex • Databricks • Elastic Stack • Docker • Ocaml • Clojure • Firebase • Go • HTML / CSS • Neo4j • Android