# Ryan Teehan

(973)-460-1672 • https://github.com/rteehas • rsteehan@gmail.com • https://www.linkedin.com/in/ryan-teehan/

**Education:** **M.S. Computer Science (2018); B.A. Mathematics (2018)**, **The University of Chicago**, **GPA: 3.66**

**Extracurricular Academic Work:**

- University of Chicago Mathematics Directed Reading Program – Fall 2015: Representation Theory of Finite Groups: Random Walks on Finite Groups. Presented the results in January of 2016
- Independent Reading: Professor Shmuel Weinberger – Summer 2018: Discussed discrepancy theory as it relates to sampling techniques

**Experience:**

- **April 2021 – Present: HuggingFace BigScience Summer of Language Models**
  - Yearlong research workshop on very large language models (LLMs) with the aim of creating and researching a multilingual LLM
  - Member of and contributor to working groups on Modeling (Architecture and Scaling), Evaluation, and Interpretability
- **January 2019 – Present: Charles River Analytics**
  **Software Engineer II (May 2020 – Present); Software Engineer I (January 2019 – May 2020)**
  - Developing probabilistic supply chain models in **Pyro** to infer the existence of missing nodes
  - Implemented a multi-resolution Bayesian time series model for real-time maintenance of the condition of complex machinery at coarse and fine time scales
  - Utilized copulas to model dependences between univariate distributions for a Monte-Carlo model that computed the relative cost of uncertainties
  - Implemented a Multi-Objective Monte Carlo Tree Search algorithm in Scala for repair schedule optimization
  - Used the **Figaro** probabilistic programming language to model satellite movements
- **September 2018 – January 2019: Infinite Analytics**
  **Data Scientist**
  - Reduced latency of **Spark** computations, implemented in Scala, by ~25%
  - Evaluated and implemented algorithms for dimensionality reduction of large (~ 300,000 x 7 million), sparse, binary matrices
  - Optimized word embeddings to improve search results for customer search engines
- **June – August 2018: Deep Skies Lab**
  **Research Assistant**
  - Generating simulations of strongly lensed galaxies to train new neural networks in a high performance computing environment using the **Lenstronomy** Python package along with **Keras**

**Technical Skills:**

- **Python**, (Pandas, Numpy, Pytorch, Tensorflow, etc.)
- **Scala**
- **Javascript** (including **Node.js**)
- **Julia**
- **SQL**
- **MATLAB**

**Projects:**

- **Multi-Word Expression Identification (2018)**
  - Worked on developing a language independent method to identify multi-word expressions
- **Backtranslations for Contranyms (2021)**
  - Explored the results of backtranslating works with contradictory meanings, especially in a legal or political context) using off the shelf translations algorithms (Presented at the AfricaNLP Workshop at EACL 2021; winner of a **Best Paper Award**)
- **Google BIG Bench (https://github.com/google/BIG-bench) (2021)**
  - Collaborated on a set of benchmark datasets for Large Language Models to the Google BIG-Bench project related to:
    - Riddles in the Kannada language
    - Humorous edits of band and music artist names
    - Inversion of normal word order and question/answer association

**Research Proposals:**

- **PRESCRIPTION SBIR Proposal – 2020**
  - Wrote and developed the probabilistic modeling component to model pharmaceutical supply chains
- **EMERGENT III BAA – 2021**
  - Writing a white paper on probabilistic supply chain modeling for a Broad Agency Announcement solicitation