Yingchen (Eric) Ma

Computational Science and Engineering Department, Georgia Institute of Technology, Atlanta, GA 30308
Email: yma473@gatech.edu | Phone: (248) 719-6226 | Homepage: https://yma17.github.io/

Education

- Ph.D. in Computer Science, **Georgia Institute of Technology** (Aug 2023 - present)
  Advisor: Srijan Kumar. *Planned minor:* Public Policy.

- M.S. in Computer Science, **Georgia Institute of Technology** (Aug 2021 - May 2023)
  Specialization: Machine Learning. *GPA:* 3.78/4.00.

- B.S. in Computer Science, **University of Michigan** (Sep 2017 - May 2021)
  GPA: 3.54/4.00.

Research Interests

Misinformation / Counter-Misinformation	Online Interventions
Social Media Analysis	Computational Social Science
Natural Language Processing	Machine Learning
Data Science	Statistics

Publications

- *Characterizing and Predicting Social Correction on Twitter* [pdf]
  Yingchen (Eric) Ma, Bing He, Nathan Subrahmanian, Srijan Kumar.
  15th ACM Web Science Conference (WebSci'23).

- *Measuring and Mitigating Group Inequalities In Resource Allocation*
  Arya Farahi, **Yingchen (Eric) Ma**, Angela Ting.
  (Submitted to) ACM Journal on Responsible Computing.

Current Projects

- **Analyzing Disparities in Social Media Counter-Misinformation:** Leading a project to investigate differences in the amount of counter-misinformation posted towards misinformation-spreading users across demographic attributes such as race, sex, and education.
  - Advised by Prof. Srijan Kumar, Assistant Professor, CSE @ Georgia Tech

Previous (completed) Projects

- **Analyzing and Predicting Twitter Counter-Misinformation:** Investigated features of misinformation-spreading tweets and users that correlate with increased likelihood to receive replies that counter / debunk said misinformation.
  - Advised by Prof. Srijan Kumar, Assistant Professor, CSE @ Georgia Tech

- **Multilingual Analysis of Autism Discussion on Twitter:** Investigated the discourse surrounding autism on Twitter across various languages, through analysis and comparison of sentiment, topical focus, and engagement statistics.
  - Course final project for CS 7650 (Natural Language) @ Georgia Tech

- **Analysis of Anxiety Discussion on Twitter:** Investigated the relationship between features of tweets that aim to seek mental health support for anxiety, and the likelihood to receive tweet engagement and supportive replies ([repository](#)).
  - Course final project for CS 6474 (Social Computing) @ Georgia Tech

- **Urban Road Repair Proposal:** Led a project to develop an end-to-end decision pipeline to propose roads in an urban setting, with the objective of maximizing economic, demographic, and accessibility related benefit to the community ([repository](#)).
  - Advised by [Prof. Arya Farahi](#), Assistant Professor, SDS @ UT Austin

- **Neural Processes Project:** Implemented Pytorch variants of neural processes (NPs), a type of deep latent probabilistic model ([repository](#)). Applied NPs to evaluate the performance of a domain adaptation algorithm proposed in a submission to ICML 2021 ([link to paper](#)).
  - Advised by [Prof. Ding Zhao](#), Assistant Professor, MechE @ Carnegie Mellon

- **Autonomous Vehicle Data Platform**: Developed a data platform in Python for storage, preprocessing, and analysis of traffic data, for research and testing of autonomous vehicles.
  - Advised by [Prof. Ding Zhao](#), Assistant Professor, MechE @ Carnegie Mellon

Work Experience

Graduate Research Assistant, **Georgia Institute of Technology**, Atlanta, GA   (Aug 2022 - present)
- Working on "Counter-Misinformation" projects - see above under "Current / Previous Projects"

Technology Solutions Intern, **Pioneer Natural Resources**, Irving, TX   (May - Aug 2023)
- Performed machine learning based risk assessment on physical, chemical, and operational factors that contribute to scaling and subsequent failure of electrical submersible pumps (ESPs) used for oil extraction.

Programming Intern, **Wal Fuel Systems**, Livonia, MI   (Dec 2021 - Apr 2022)
- Implemented a genetic algorithm to propose more distance, time, and fuel efficient company truck delivery schedules ([repository](#)).
- Delivered end-to-end pipeline for company employees to run, that utilizes the algorithm to generate daily truck routing decisions based on updatable location and demand data.

Skills

- **Technical (computational):** Machine learning, deep learning, natural language processing, graph learning, applied data analysis, data structures, algorithms.
- **Technical (mathematical):** Optimization, linear algebra, probability, statistics, calculus.
- **Programming languages:** Python (proficient), C++ (proficient), SQL, C, Java.
- **Libraries/frameworks/tools**: ML libraries (Pytorch, Tensorflow, scikit-learn), Numpy, Pandas, Matplotlib, Plotly, Streamlit, Jupyter Notebook, Google Colab, Git, Linux, Docker.