SIYI WU

Email: reyna.wu@mail.utoronto.ca
Mobile: 437-599-5541

EDUCATION

University of Toronto Sep 2022 - Dec 2027 (Expected)
Doctor of Philosophy in Computer Science; Supervisor: Steve Easterbrook

Hong Kong University of Technology Sep 2021 - June 2022
Master of Science in Financial Technology

Wake Forest University Sep 2017 - May 2021
Bachelor of Science in Computer Science & Mathematical Statistics

- **Academics: Honors**, GPA:3.97/4.00, **Summa Cum Laude**, Dean's List, **Upsilon Pi Epsilon**, **Phi Beta Kappa**, **Pi Mu Epsilon**, **Alpha Phi Omega**
- **Awards:** URECA research fellowship ($8000), Grace Hopper Scholarship 2020.

RESEARCH INTERESTS

Climate Informatics; Machine Learning; Computational Sustainability; Trustworthy & Explainable AI; Human-Computer Interaction; Environmental Justice

RESEARCH PROJECTS

- **Siyi Wu**, Steve M Easterbrook (2023), Regional Studies of Multimodel Ensemble of Climate Projections for Enhanced Interpretability and Performance using Machine Learning Approaches, Abstract presented at AGU23, 11-15 Dec.

  We apply machine learning methods to estimate temperature and precipitation with 25 CMIP6 models, and provide interpretations of ML results in terms of contribution of each CMIP6 model in estimating with regional analyses

- **Siyi Wu**, Steve M Easterbrook (2023), Improving Interpretability and Performance in Multi-Model Ensemble CMIP6 Climate Projections through Machine Learning, Abstract presented at AGU23, 11-15 Dec.

  We propose a framework to explore the information offered by the multi-model ensemble simulations with in CMIP6, which aims to utilize the strengths of ML in MME construction while preserving interpretability and explainability of MME. Our findings highlight the strengths and limitations of individual GCMs in predicting maximum temperature.

- **Siyi Wu**, Steve M Easterbrook, Ishtiaque Ahmed (2023). "Ethical Aspects of Artificial Intelligence in Environmental Justice". Workshop on HCI for Climate Change at CHI Conference on Human Factors in Computing Systems, 2023.

  We use machine learning methods to predict the presence of UOG wells or flares based on sociodemographic features, aiming to discern whether marginalized groups bear a disproportionate burden of environmental harms; and discussed limitations of AI models and the need for considering ethical dimensions of AI in environmental justice problems, contributing to an deepen understanding of critical computing and environmental injustice

- Han Qiao, **Siyi Wu**, Christoph Becker, "Near Data" and "Far Data" for Urban Sustainability: How Do Community Advocates Envision Data Intermediaries? Under Review in ACM CHI Conference on Human Factors in Computing Systems(ACM CHI 2024).

  Through interviewing 17 advocates, we identified four design visions for data intermediary tools, exploring the tensions inherent in each vision and discussing how data feminism principles could guide data intermediaries in navigating these tensions

- Eshta Bhardwaj, Harshit Gujral, **Siyi Wu**, Ciara Zogheib, Tegan Maharaj, Christoph Becker, Machine Learning Data Practices through a Data Curation Lens: An Evaluation Framework. Under Review in Conference on fairness, accountability and transparency (ACM FAcct'24).

  We examine data practices in machine learning (ML) dataset development through the perspective of data curation, developing a framework to evaluate ML datasets using data curation principles. It reveals challenges in applying data curation concepts to ML, including difficulties in interpretation, the depth of expertise required, and documentation scope, and proposes a comprehensive framework to integrate data curation methods into ML data practices.

RESEARCH EXPERIENCE

Toronto Climate Observatory TCO   Sep 2022 - Present
Graduate Researcher (Supervised by Prof. Steve Easterbrook)
 investigating the intersection of environmental justice, specifically exploring urban heat island effects, particularly examining the role of green spaces and cooling centers in relation to socioeconomic factors in urban areas

The DataMine Research Group DataMine@WFU   Jan 2020 - May 2021
 Honor Thesis: dimensionality reduction of single cell RNA sequencing(scRNA-seq) data
 Granted 2020 Summer Wake Forest Research Fellowship; Reduced dimensionality of sparse and large scRNA-seq datasets for tumor detection using PCA, t-SNE, UMAP and GAN on 11 public datasets; Identified the best dimensionality reduction method, UMAP, of scRNA-seq data in terms of best NMI score from clustering analysis

Wake Forest Research Fellowship 2019 Supervised by Robert Erhardt   Jan 2019 - May 2021
 Honor Thesis: Climate Model Compression and Predictability
 Granted 2019 Summer Wake Forest Research Fellowship; Quantify the predictability of compressed climate data; Implemented empirical orthogonal functions (EOF) to compress datasets and visualize the patterns, predicted temperature and precipitation using time series models, and evaluated results via root means squared error; Proved possibilities of summarizing climate data with EOF and giving better prediction with VAR models

TEACHING EXPERIENCE

Teaching Assistant, UofT, Introduction to Databases (CSC343)   Winter 2024
Teaching Assistant, UofT, Computers and Society(CSC300)   Winter 2024
Teaching Assistant, UofT, Introduction to Databases (CSC343)   Fall 2023
Teaching Assistant, UofT, Introduction to Computer Science (CSC148)   Winter 2023
Teaching Assistant, Neoscholar CIS, Machine Learning: Theory and Practical Algorithms by Victor Adamchik 2021
Teaching Assistant, Neoscholar CIS, Computational Biology by Sorin Istrail   2021
Teaching Assistant, Neoscholar CIS, Planning for Autonomous robots by Nick Hawes 2021
Teaching Assistant, Neoscholar CIS, Algorithms by David Woodruff   Summer, Spring, Winter 2020
Teaching Assistant, Wake Forest University, Introduction to Computer Science (CSC111)   Winter 2020

SKILLS

Programming Languages: Python, R, C, C++, Javascript, HTML, MySQL
Frameworks and Tools: Keras, Scikit-learn, PyTorch, Pandas, NumPy, D3, Matplotlib, LaTeX