Ali Turfah

Ann Arbor, MI 48105, USA
aturfah@umich.edu
aturfah.github.io

Education

Ph.D. in Biostatistics, *University of Michigan* Sept. 2021 – Present
Adviser: Xiaoquan (William) Wen

M.A. in Statistics, *Columbia University* Sept. 2019 – Dec. 2020

B.S.E. in Data Science, *University of Michigan* Sept. 2014 – April 2018
Minor in Mathematics

Research Experience

Improving COVID-19 Vaccine Safety Analysis, *University of Michigan* June 2021 – March 2023
Improve estimation for association of rare adverse events with COVID-19 vaccines based on VAERS
- Implement Bayesian logistic regression model to conduct vaccine safety analysis using Dirichlet process mixture (DPM) prior to improve estimates via information sharing
- Incorporated negative control and adverse event group enrichment procedures into MCMC sampler
- Demonstrated drastic improvements in MSE and DIC for the DPM model compared to models with independent priors in simulation studies
- Identified 17 adverse events and 5 groups significantly associated with COVID-19 mRNA vaccines, confirming findings in medical literature
- Accepted as poster to ICSA 2023 conference; manuscript under review

Effect of Diabetes on Mouse Cognitive Function, *University of Michigan* Sept. 2021 – April 2022
Understand the relationship between diet and performance on cognitive tasks in diabetic mice
- Implement piece-wise longitudinal model to flexibly capture the relationship between diet and time to complete the Morris Water Maze
- Conduct survival analysis to fit the effect of diet on mice ability to escape from Puzzle Box arenas
- Compare the predictive ability of mouse weight and glucose tolerance test result on puzzle box escape times

Medical Evidence Normalization, *Columbia University* Jan. 2021 – June 2021
Develop standard representation of findings in medical literature to facilitate evidence computation tasks
- Designed and implemented framework for the normalization of PICO Observation elements for the extraction and representation of medical evidence from medical literature
- Built preliminary system to detect contradictions in the findings of medical literature, scaling process to be applied to 100,000 abstracts published on PubMed from 2015-2020
- Accepted in MedINFO

Medical Evidence Dependency-informed Attention, *Columbia University* April 2021 – Sept. 2021
Incorporate medical evidence relations based on the PICO framework into neural attention model
- Developed interpretable novel attention mechanism incorporating evidence from medical literature, representing medical evidence by dependency relations based on PICO elements
- Benchmarked performance against standard attention in BioBERT, improving performance by up to +0.3 F1 on medical question-answer datasets
- Published in JAMIA

Automating Drug Safety Reporting, *University of Michigan* Jan. 2017 – Dec. 2017
Automate detection of reportable adverse events in drug reports to FDA via NLP models
- Researched natural language processing literature to determine appropriate algorithms; final model used GloVe embedding with an LSTM architecture to determine if an article's findings should be reported to the FDA
- Designed data normalization process to simplify training and streamline implementation of production system
- Implemented final algorithm to perform document classification, achieving 80% precision and 96% recall
- Organized meetings with team members, company sponsors, and faculty advisers to report and present progress

Industry Experience

Data Analyst Intern, *Icahn School of Medicine at Mount Sinai*, New York, NY Feb. 2020 – Aug. 2020
Student position in Dr. Roxana Mehran's Interventional Cardiology group reporting to Dr. Samantha Sartori
- Generated analysis and visualizations for cardiovascular outcomes studies
- Set up the automated generation of reports for large-scale studies for presentation to sponsors and review boards

Data Scientist, *Algo*, Troy, MI June 2018 – July 2019

- Lead design and implementation of new forecasting modules, improving accuracy by 12% over original forecaster
- Appointed customer-facing technical lead of behind-schedule demand planning project, successfully completing the project on time by incorporating new development and requirements-gathering practices
- Optimized and corrected SQL in daily ETL processes, lessening system strain and reducing processing errors
- Mentored new employees through on-boarding process with pair programming exercises, overviews of relevant statistical concepts, and one-on-one meetings to discuss progress

Machine Learning Engineering Intern, *Algo*, Troy, MI May 2017 – Aug. 2017

- Integrated statistical clustering methods into API, enabling on-demand analysis of store performance data
- Developed proof-of-concept system to automate unit testing, simplifying deployment and reducing bug reports
- Wrote web scraper to augment available customer data using information from Rotten Tomatoes and IMDB, retrieving movie meta data information including genre, ESRB rating, box office, director, and audience response
- Built process to export customer data in email and powerpoint formats, allowing results to be readily shared

Publications

Turfah, **Ali**, Xiaoquan Wen, and Lili Zhao (2023). "Non-parametric Bayesian mixture model to study adverse events of COVID-19 vaccines". In: *arXiv preprint arXiv:2306.02123*.

Miller, Erin, Rima Charara, **Ali Turfah**, and Jennifer Mendez (2022). "Optimizing older adult blood pressure screening in a community setting by interprofessional students". In: *Innovation in Aging* 6.Suppl 1, p. 517.

Turfah, **Ali**, Hao Liu, Latoya A Stewart, Tian Kang, and Chunhua Weng (2022). "Extending PICO with Observation Normalization for Evidence Computing". In: *Studies in health technology and informatics* 290, pp. 268–272.

Chiarito, M., A. Roumeliotis, D. Cao, D. Power, S. Sartori, Z. Zhang, A. Reisman, T. Mtisi, M. Nardin, J. Nicolas, H. Qiu, **A. Turfah**, et al. (2021). "Prevalence and prognostic impact of high bleeding risk status in patients undergoing percutaneous coronary intervention for left main coronary artery disease". In: *Journal of the American College of Cardiology* 77.18 Supplement 1, pp. 1148–1148.

Kang, Tian, **Ali Turfah**, Jaehyun Kim, Adler Perotte, and Chunhua Weng (2021). "A neuro-symbolic method for understanding free-text medical evidence". In: *Journal of the American Medical Informatics Association*.

Nardin, M., D. Cao, M. Chiarito, J. Nicolas, S. Sartori, Z. Zhang, H. Qiu, **A. Turfah**, G. Giustino, R. Chandiramani, et al. (2021). "Prognostic value of the academic research consortium for high bleeding risk criteria in patients with diabetes melitus undergoing PCI". In: *Journal of the American College of Cardiology* 77.18 Supplement 1, pp. 167–167.

Nicolas, J., B. Claessen, D. Cao, M. Chiarito, S. Sartori, H. Qiu, R. Goel, M. Nardin, A. Roumeliotis, B. Vogel, **A. Turfah**, et al. (2021). "A Sex Paradox in Clinical Outcomes Following Complex Percutaneous Coronary Intervention". In: *International journal of cardiology* 329, pp. 67–73.

Nicolas, J., D. Cao, B. Claessen, S. Sartori, R. Chandiramani, A. Roumeliotis, R. Goel, A. Camaj, F. Beerkens, **A. Turfah**, G. Dangas, U. Baber, S. Sharma, A. Kini, and R. Mehran (2020). "Long-term outcomes in high-bleeding risk patients undergoing PCI for acute coronary syndromes: results from a large single-center PCI registry". In: *European Heart Journal* 41.

Nicolas, J., D. Cao, B. Claessen, S. Sartori, A. Roumeliotis, R. Goel, R. Chandiramani, G. Stefanini, **A. Turfah**, S. Chen, G. Dangas, U. Baber, S. Sharma, A. Kini, and R. Mehran (2020). "Intersection of the Academic Research Consortium - high bleeding risk criteria in patients undergoing PCI for acute coronary syndromes: insights from a high-volume single center registry". In: *European Heart Journal* 41. (Under Review).

Nicolas, J., B. E Claessen, D. Cao, M. Chiarito, S. Sartori, H. Qiu, R. Goel, M. Nardin, A. Roumeliotis, B. Vogel, **A. Turfah**, R. Chandiramani, U. Baber, N. Barman, J. Sweeny, P. Krishnan, A. Kini, S. K. Sharma, G. D. Dangas, and R. Mehran (2020). "Sex Disparities Among Patients Undergoing Complex Percutaneous Coronary Intervention (PCI): Insights From a Single-Center Large-Volume PCI Registry". In: *Journal of the American College of Cardiology* 76.17 Supplement S, B162–B163.

Extra-Curricular Activities

Alumni Mentor, *Columbia University* Jan. 2023 – Present

Mentor current Statistics M.A. students in preparation for Ph.D. programs

Peer Mentoring Committee, *University of Michigan* June 2022 – June 2023
Support to incoming students by mentoring programs and communicate concerns to the department

Graduate Student Mentor, *Ardsley High School*, New York, NY Nov. 2020 – March 2021
Assist students with statistical analysis components of Westchester Science and Engineering Fair projects

Website Developer, *Columbia University* Oct. 2019 – March 2020
Developed ASA Statistical Learning/Data Science section's 2020 conference website

Mathematics Tutor 2012 – 2019
Tutoring for college-level math courses and math sections of GRE and PCAT exams

President, *Michigan Biological Software Team*, University of Michigan Oct. 2015 – April 2018
- Gold Medalist at International Genetically Engineered Machine (iGEM) Competition 2016, 2017
- Best Software Award at iGEM Competiton 2016

Workshop Leader, *Canton Public Library*, Canton, MI Jan. 2016 – March 2016
Organized and delivered superhero themed chemistry workshop to elementary school library patrons

Webmaster, *Michigan Biological Software Team*, Ann Arbor, MI Jan. 2015 – Aug. 2015
Bronze Medalist at iGEM Competition 2015