Ali Turfah

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

Education

September 2021 to Present
Ph.D. in Biostatistics, *University of Michigan*, Ann Arbor, MI, USA.
- GPA: 3.90

September 2019 to December 2020
M.A. in Statistics, *Columbia University*, New York, NY, USA.
- GPA: 3.90
- Awards and Honors:
  - Chair's List of Academic Achievement: Spring 2020
  - First Place: DataFest Fall 2019

September 2014 to April 2018
B.S.E. in Data Science, *University of Michigan*, Ann Arbor, MI, USA.
- GPA: 3.31 *Cum Laude*
- Minor in Mathematics
- Awards and Honors:
  - Dean's Honor List: Fall 2017, Winter 2018
  - University Honors: Fall 2017, Winter 2018
  - Best Software Award: iGEM Competition 2016
  - Gold Medalist: iGEM Competition 2016, 2017

Core Competencies

- **Programming Languages**: Python, R, SQL, Javascript
- **Skillsets**: Data Mining, Machine Learning, API Development, Project Management
- **Technologies**: Numpy, Tensorflow, BUGS/JAGS, PLINK, Chart.js

Experience

July 2021 to Present
Research Assistant, *University of Michigan Department of Biostatistics*, Ann Arbor, MI, USA.
- Developed Bayesian model to study the difference in the prevalence of adverse events between the COVID-19 vaccine and other vaccines, as well as compare these differences between the various COVID vaccines
- Expanded model to group the effects based on a Dirichlet clustering approach to address convergence issues, as well as incorporated negative controls to assess bia

April 2020 to July 2021
Research Assistant, *Columbia University Department of Biomedical Informatics*, New York, NY, USA.
- Developed novel attention mechanism that is interpretable and incorporates evidence from medical literature, improving performance over standard attention in BioBERT by up to +0.3 F1 on medical QA datasets
- 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

February 2020 to August 2020
Data Analyst Intern, *Icahn School of Medicine at Mount Sinai*, New York, NY, USA.
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

June 2018 to July 2019
Data Scientist, *Algo*, Troy, MI, USA.
- 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

May 2018 to January 2019
Student Researcher, *University of Michigan*, Ann Arbor, MI.
Independent research under Dr. Yves Atchade to develop multinomial bayesian models for cancer-type classification.
- Derived and implemented preliminary classification model, including optimizations for computational efficiency
- Enhanced original model using spike-slab priors to account for sparsity, in addition to Polya-Gamma augmentation to yield a closed form posterior for binary case

September 2017 to April 2018 **Research Assistant**, *Center for Molecular Imaging*, Ann Arbor, MI, USA.
Student position to support Dr. Craig Galban and his team in medical imaging data analysis.
- Designed and implemented application to automate processing of new data sets, saving time and reducing errors
- Built tool consolidating all image analysis scripts into single interface, streamlining process flow
- Assisted in the preparation of incoming datasets, including orienting, rotating, and cropping of images

May 2017 to August 2017 **Machine Learning Engineering Intern**, *Algo*, Troy, MI, USA.
- 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

January 2017 to December 2017 **Student Researcher**, *ProQuest*, Ann Arbor, MI, USA.
Student-run pilot research project to improve the pharmaceutical literature review process using machine learning.
- 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

Publications

[1] Ali Turfah, Hao Liu, Latoya A Steward, Tian Kang, and Chunhua Weng. Extending pico with observation normalization for evidence computing. *MedInfo*, 2021.

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

[3] J. Nicolas, B. 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. Sharma, G. Dangas, and R. Mehran. A Sex Paradox in Clinical Outcomes Following Complex Percutaneous Coronary Intervention. *International journal of cardiology*, 329:67–73, 2021.

[4] J. Nicolas, 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. 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. *European Heart Journal*, 41, 2020. (Accepted as abstract).

[5] J. Nicolas, 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. Long-term outcomes in high-bleeding risk patients undergoing PCI for acute coronary syndromes: results from a large single-center PCI registry. *European Heart Journal*, 41, 2020. (Accepted as abstract).

[6] J. Nicolas, 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. Sex Disparities Among Patients Undergoing Complex Percutaneous Coronary Intervention (PCI): Insights From a Single-Center Large-Volume PCI Registry. *Journal of the American College of Cardiology*, 76(17 Supplement S):B162–B163, 2020. (Accepted as abstract).

Extra-Curricular Activities

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

October 2019 to March 2020 **Website Developer**, *Columbia Department of Statistics*, New York, NY, USA.
- Developed ASA Statistical Learning/Data Science section's 2020 conference website

October 2015 to April 2018 **President**, *Michigan Biological Software Team*, Ann Arbor, MI, USA.
- Gold Medalist at International Genetically Engineered Machine (iGEM) Competition 2016, 2017
- Best Software Award at iGEM Competiton 2016

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

January 2015 to August 2015 **Webmaster**, *Michigan Biological Software Team*, Ann Arbor, MI, USA.
- Bronze Medalist at iGEM Competition 2015