Michael Hellstern

mikeh1@uw.edu | mikehellstern.github.io | mikehellstern | Michael Hellstern | Michael Hellstern

Summary of Qualifications

- **Experience**: Ph.D. statistician with over 7 years of technical experience in the development and application of novel AI and deep learning methodologies and 3 years of industry experience in analytics consulting.
- **Consulting and Collaboration**: Proven ability to bridge the gap between technical methodologies and business value, with specific expertise in managing client relationships and communicating complex concepts to non-expert stakeholders.
- **Technical Lead**: Proven track record of leading end-to-end projects, from problem identification to scalable, high-performance implementation in Python (PyTorch), R, and C++.
- **Technical Skills**: Advanced knowledge of Python and PyTorch for machine learning and GPU acceleration, LangChain, Ollama, vLLM, LLM Foundation Models, SQL, R, C++, SAS.

Education

Ph.D. in Biostatistics, University of Washington — *Seattle, WA*
Thesis: Methods for Time Series Network Analyses (advised by Ali Shojaie) — *Sep 2020 - Aug 2025*

M.S. in Statistics, University of Washington — *Seattle, WA*
Jun 2018 - Aug 2020

B.A. in Economics, Williams College — *Williamstown, MA*
Sep 2011 - Jun 2015

Experience

University of Washington — *Seattle, WA*

Postdoctoral Research — *Sep 2025 - Present*

- Developing a multi-stage LLM framework, including LLM-as-a-judge, for causal inference and precision medicine using EHR data.

Doctoral Research — *Jun 2018 - Sep 2025*

- Initiated, conceptualized, and developed an end-to-end, model-agnostic, framework to jointly learn AI and deep learning models and detect changes in model performance. Built scalable PyTorch pipeline, including processing of simulation and COVID-19 data. Leveraged GPU acceleration on a HPC.
- Improved forecast accuracy in predicting COVID-19 outcomes by over 20% by developing a new order selection method in Vector Autoregressive models.
- Developed a computationally efficient inference procedure for high-dimensional networks enabling real-time analysis. Used to analyze dynamic brain networks under external stimulation aiding in treatment of neurodegenerative disorders.

Statistical Consulting — *Fall 2022*

- Provided guidance on statistical analysis and programming for clients across diverse domains including pediatrics, nutrition sciences, and forestry to translate real-world research questions into statistical and analytical solutions, explained in clear, accessible terms.

Research Assistant — *Jun 2018 - Jun 2019*

- Developed and maintain the `netgsa` R package for network-based gene set analysis. Achieved major improvements in computational efficiency - up to 40x faster - through unsupervised clustering and C++ integration.

Analysis Group — *Boston, MA*

Senior Analyst — *Jul 2015 - May 2018*

- Collaborated with external clients on analytical projects, translating large-scale insurance claims and EMR data into actionable business insights that resulted in 10 peer-reviewed publications. Conducted analyses in SQL, SAS, and R.
- Member of experimentation platform team designing core statistical tools and interactive user interface in R and RShiny. Integrated feedback from case teams and clients. Platform was used across dozens of case teams to streamline analysis time from days to hours.

Publications / In preparation

[1] **Hellstern M**, Shojaie A. "Dynamic Deep Learning for Detecting Changes in Biomedical Processes". *Under review*.

[2] **Hellstern M**, Shojaie A. "Order Selection in Vector Autoregression by Mean Square Information Criteria". DOI: 10.48550/arXiv.2511.19761

[3] **Hellstern M**, Kim B, Shojaie A. "Assumption-Lean Inference for Spectral Differential Network Analysis for High-dimensional Time Series". In preparation. *Draft available on request.*

* Received a *Best Student Paper Award* from the ASA Statistical Learning and Data Science (SLDS) section. Presented at JSM 2025.

[4] **Hellstern M**, Kim B, Harchaoui Z, Shojaie A. "Spectral Differential Network Analysis for High-dimensional Time Series". *AISTATS (2025)*. DOI: 10.48550/arXiv.2412.07905

* Received a *Best Student Poster Honorable Mention* from the ASA Statistical Learning and Data Science (SLDS) Section.

[5] **Hellstern M**, Ma J, Yue K, Shojaie A. "netgsa: Fast computation and interactive visualization for topology-based pathway enrichment analysis". *PLOS Computational Biology 17.6 (2021): e1008979*. DOI: 10.1371/journal.pcbi.1008979

[6] Desai U, Kirson NY, Kim J, Khunti KK, King SB, Trieschman E, **Hellstern M**, Hunt PR, Mukherjee J. "Time to Treatment Intensification After Monotherapy Failure and Its Association With Subsequent Glycemic Control Among 93,515 Patients With Type 2 Diabetes". *Diabetes care 41.10 (2018): 2096-2104*. DOI: 10.2337/dc17-0662

[7] DerSarkissian M, Lefebvre P, Joshi K, Brown B, Lafeuille MH, Bhak R, **Hellstern M**, Bobbili P, Shiner B, El Khoury A, Young-Xu Y. "Health Care Resource Utilization and Costs Associated with Transitioning to 3-month Paliperidone Palmitate Among US Veterans". *Clinical Therapeutics 40.9 (2018): 1496-1508*. DOI: 10.1016/j.clinthera.2018.07.011

[8] Fuld A, Young-Xu Y, Li S, Pilon D, Bhak R, Duchesneau E, **Hellstern M**, Kamstra R, Behl A, Lefebvre P, Freedland S. "Predictors of overall survival (OS) in veterans with non-metastatic castration resistant prostate cancer (nmCRPC)". *Journal of Clinical Oncology 36 (2018): e17057-e17057*. DOI: 10.1200/JCO.2018.36.15_suppl.e17057

Honors & Awards

- *Best Student Paper Award* from the ASA Statistical Learning and Data Science (SLDS) *2025*
- *Best Student Poster Honorable Mention* from the ASA SLDS *2024*
- *Donovan J. Thompson Award* for best combined performance on PhD qualifying exams in UW Biostatistics *2021*

Invited Talks

Assumption-Lean Inference for Spectral Differential Network Analysis of High-Dimensional Time Series

- Joint Statistical Meetings (JSM) *2025*
- Western North American Region of the International Biometric Society (WNAR) *2025*

Spectral Differential Network Analysis for High-Dimensional Time Series

- Western North American Region of the International Biometric Society (WNAR) *2024*
- Joint Conference on Computational and Financial Econometrics and Computational and Methodological Statistics (CFE-CMStatistics) *2024*

Service

- Reviewer for Journal of Machine Learning Research, Journal of the Royal Statistical Society: Series B, Journal of the American Statistical Association

Interests

- Mountain biking, whitewater kayaking, trail running, skiing, ski touring