Yujie (Jeffrey) Jiang

+1 (336) 829 2490 • yujiejiang679@gmail.com • Google Scholar • in yujiejiang

Education

Rice University **Houston, TX**
Ph.D. in Statistics, GPA: 3.93/4.0 *Aug 2018 – May 2024 (Expected)*
- American Statistical Association (ASA) Student Paper Award (2021)
- Ken Kennedy Institute Computational Science and Engineering Fellowship (2018)

Wake Forest University **Winston-Salem, NC**
B.S. in Computer Science and B.S in Mathematical Statistics, with honor *Aug 2014 – May 2018*
- Graduation Honor: Magna Cum Laude
- Atlantic Coast Conference Inter-Institutional Academic Collaborative (ACCIAC) Fellowship (2017)
- Pi Mu Epsilon honorary society & Upsilon Pi Epsilon honorary society (2017)
- Wake Forest Research Fellowship (2016)
- Dean's List every semester
- Second violin in Wake Forest University Symphony Orchestra

Skills

Programming Languages: Python, R, MATLAB, C/C++, SQL, JAVA, JavaScript, HTML

Framework & Tools: Git, Pandas, Numpy, Scikit-Learn, TensorFlow, LaTeX, Linux, BioRender, Bioconductor

Research Focuses: Optimization, High-Dimensional Data Analysis, Optimization, Machine Learning, Tensor Decomposition

Professional Experience

Sanofi U.S **Bridgewater, NJ**
Statistician Intern *May 2023 – Aug 2023*
- Built a Bayesian analysis software that evaluates statistical power of clinical trial outcomes, estimates assurance in proposed trials, and optimizes phase 2 & 3 study designs, thereby facilitating decision-making process in drug development cycle.
- Designed comprehensive testing protocols and simulated real-world scenarios to ensure software accuracy and reliability.
- Wrote core functions in R that contribute to the software's feature development and drafted materials for dissemination.

Quantlab Financial, LLC **Houston, TX**
Quantitative Research Intern *Jun 2020 – Aug 2020*
- Implemented decision trees and long short-term memory (LSTM) neural network models in alpha research, leveraging vast trading data to drive insights and facilitate data-driven decision-making.
- Experimented pairs trading strategies with cointegration and stochastic-based optimal entry/exit model, back-tested those signals and demonstrated 3 out of 21 with out-of-sample Sharpe ratio > 3 and annual return > 30%.
- Created and optimized core functions that extract valuable insights and perform statistical analysis on trading records using a proprietary programming language.

Wake Forest Demon Deacons Women's Basketball **Winston-Salem, NC**
Basketball Statistical Analyst *Sep 2016 – Sep 2017*
- Assisted in designing data strategy and supporting in-game analysis as well as player development.
- Restructured and maintained the statistical database for the team; constructed data analysis and calculated advanced metrics.
- Generated scouting reports to help game preparation and presented statistical analysis before each game.

Research Experience

Statistical Bioinformatics Lab, University of Texas MD Anderson Cancer Center **Houston, TX**
Research Assistant *Jan 2020 – present*
- Developing and applying computational methods to study the evolution of the human genome as well as the cancer genome.
- Proposing new methodologies and corresponding algorithms for large-scale complex data analysis on subclonal reconstruction.
- Advancing methodology ideas and efficient software that accurately analyze high-throughput genomic data.

Genevera Allen's Statistical Learning Group, Rice University **Houston, TX**
Research Assistant *Jun 2018 – Dec 2019*
- Constructed a new extension of the decomposition model that is specifically focused on modeling semi-symmetric tensors, and have them implemented to MRI & neuroimaging studies.

Ballard Research Group on Tensor Analysis, Wake Forest University **Winston-Salem, NC**
Research Assistant *Sep 2016 – May 2018*
o Conducted exploratory research on high-performance tensor computations, delivering weekly presentations with a specific focus on low-rank approximation topics.
o Developed efficient algorithms for dense canonical polyadic tensor decomposition, and implemented in both C and MATLAB.

Other Projects

Biological Data Analysis: Annotation and Interpretation of High-throughput Sequencing Data *Fall 2022*
o Analyzed single-cell RNAseq data of wild-type & chronic kidney disease mice to identify potential cellular & molecular etiology.
o Investigated whole exome sequencing data from 100 sudden cardiac death families to identify *de novo* diseases-causing variants.
o Tracked the evolution trajectory from papillary thyroid cancer to anaplastic thyroid cancer using subclonal reconstruction and mutation timing technique, on over 550 whole genome and whole exome sequencing samples.

Tensor Analysis: Supervised Tensor Decomposition for Semi-symmetric Tensors *Summer 2019*
o Explored the underlying relationship among PCA, CCA, and PLS, especially on high order spaces.
o Proposed general optimization formulation that expresses both tensor CCA and tensor PLS methods.
o Present one tensor CCA algorithm and two tensor PLS algorithms for modeling semi-symmetric tensors in real-world problems.

Web Development: Bike Sharing at Wake Forest University *Spring 2017*
o Led a team of 4 members to initiate and deliver a full-stack web service for the university's bike sharing program.
o Built features including registration, database management, and bike reservation & return by JavaScript, HTML, and SQL.
o Engineered under test-driven methodologies, and kept regular communication with clients to ensure user needs were fulfilled.

Publications

[1] C. Zheng, Y. Wei, . . . , **Y. Jiang** *et al.*, "Multiomics analyses reveal DARS1-AS1/YBX1-controlled post-transcriptional circuits promoting glioblastoma tumorigenesis/radioresistance," *Science Advance*, 2023.

[2] S. Guo, Y. Dong, C. Wang, **Y. Jiang**, R. Xiang, L. Liu, H. Luo, and L. Fan, "Integrative analysis reveals the recurrent genetic etiologies in idiopathic pulmonary fibrosis," *QJM: An International Journal of Medicine*, 2023.

[3] S. Guo, Y. Dong, R. Du, j. Liu, H. Xu, Y. Liu, **Y. Jiang**, H. Huang, L. Fan, and R. Xiang, "Single-cell analysis reveals the cellular and molecular landscape of rtn3-null-induced renal fibrosis," *Under review*, 2023.

[4] Y. Dong, Z. Lin, J. Yun, Y. Zhi, L. Zhu, Y. Sun, **Y. Jiang**, S. Guo, Y. Li, and L. Fan, "Integrated analysis of wes and scrna-seq data provides new insights into unexplained recurrent pregnancy loss in 13 chinese families," *Under review*, 2023.

[5] A. Maniakas, Y. C. Henderson, H. Hei, S. Peng, Y. Chen, **Y. Jiang**, S. Ji, M. Cardenas *et al.*, "Novel anaplastic thyroid cancer PDXs and cell lines: Expanding preclinical models of genetic diversity," *The Journal of Clinical Endocrinology & Metabolism*, 2021.

[6] **Y. Jiang**, K. Yu, S. Ji, S. J. Shin, S. Cao, M. D. Montierth, L. Huang, S. Kopetz, P. Msaouel *et al.*, "CliP: subclonal architecture reconstruction of cancer cells in dna sequencing data using a penalized likelihood model," *bioRxiv*, 2021.

[7] K. Hayashi, G. Ballard, **Y. Jiang**, and M. J. Tobia, "Shared-memory parallelization of MTTKRP for dense tensors," in *Proceedings of the 23rd ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming*, 2018.

[8] **Y. Jiang** and R. Plemmons, "Blind hyperspectral unmixing through low-rank matrix and tensor approximation," *Submitted as the honor thesis to the Department of Mathematics & Statistics, Wake Forest University*, 2018.

[9] **Y. Jiang**, K. Yu, M. D. Montierth, S. Ji, S. J. Shin, S. Guo, S. Cao, S. Kopetz, P. Msaouel *et al.*, "Pan-cancer analysis of 7,711 cancers in the extent and impact of intratumor heterogeneity using a regularized likelihood model," *Under preparation*, 2023.

[10] S. Guo, X. Cheng, **Y. Jiang**, X. Liu, S. Ji *et al.*, "DeMixSC: a pseudo-bulk based gene selection method for bulk transcriptomics deconvolution," *Under preparation*, 2023.

Selected Presentations

o Tracking pan-cancer genomic evolution using a regularized likelihood model, *ENAR, 2023*
o Tracking cancer evolution with whole-genome and whole-exome sequencing using reconstructed subclonal architecture, *26th International Conference on Research in Computational Molecular Biology (RECOMB), 2022*
o CliP: A model-based method for subclonal architecture reconstruction using regularized maximum likelihood estimation, *Joint Statistical Meeting (JSM), 2021*
o CliP: A model-based method for subclonal architecture reconstruction using regularized maximum likelihood estimation, *American Society of Human Genetics (ASHG) Annual Meeting, 2020*
o Algorithmic Optimization for Dense Tensor Decomposition, *Wake Forest Undergraduate Research Day, 2017*
o Binary Sequences with Structural Delays, *Mathematical Association of America MathFest Annual Conference, 2016*