Tara Friedrich | Resume

tfriedrich.solutions@gmail.com
Open to opportunities in the SF Bay Area or remote

Programming Languages: Python, R, SQL

Data Science: Pandas, Numpy, Scikit-learn, PyTorch, Matplotlib, Seaborn, Ggplot2, Bioconductor, Keras, RShiny

Computational Resources: High Performance Cluster, AWS (S3, EC2, Batch, Lambda), Google Cloud Compute, Kubernetes, Docker, WDL, Nextflow, Git, JIRA, SQL (MySQL, Spark), noSQL (MongoDB, Amazon DynamoDB)

Statistics and Machine Learning: Hypothesis Testing, Linear and Generalized Linear Modeling, Multivariate Modeling, Clustering, Dimensionality Reduction, Supervised and Unsupervised Classification, Random Forests, Model Selection

Summary

Bioinformatics Data Scientist with expertise in multimodal clinical and genetic data analysis. Proficient in advanced statistical methods and converting ideas into code that bridge the gap between domain experts and engineering teams. Skilled in delivering actionable insights to drive innovation in healthcare and research.

Experience

Senior Bioinformatics Data Scientist - Natera — Feb 2022–August 2024

· Partnered with lab scientists to enhance RNA-seq assays, focusing on detecting oncogenic fusions in noisy FFPE tumor samples, leading to improved assay sensitivity and specificity.

· Developed custom algorithms to identify Illumina reads mapping to fusion junctions to explain discordance between fusion callers (*Pizzly, Arriba, etc.*) .

· Collaborated with clinical genomic scientists and engineers to develop software for variant interpretation.

· Integrated splicing prediction models, population statistics, structural and functional information into a classifier.

Bioinformatics Data Scientist - Fabric Genomics — Nov 2019–Jan 2022

· Optimized the ACE product algorithm, a production-level tool for prioritizing clinically pathogenic genetic variants, by implementing regression and unit testing.

· Responsible for ensuring the accuracy of data and algorithm logic while adhering to ACMG clinical guidelines.

· Enhanced variant effect prediction classifiers to increase precision in identifying likely pathogenic variants and flagging borderline cases for use by CAP/CLIA regulated labs using statistics and machine learning.

· Leveraged distributed computing frameworks like Apache Spark to scale data classification workflows and optimize ETL processes, streamlining data analysis pipelines for large-scale datasets.

Bioinformatics Programmer - UCSF — July 2018–July 2019

· Identified and prioritized candidate SNPs from whole exome sequencing data for CRISPR target screens.

· Conducted exploratory analysis on single-cell RNA-seq liver data (Tabula Muris project) using clustering, visualization, and differential expression methods to identify key cell populations.

· Designed and analyzed high dimensional datasets derived from epigenetic reprogramming experiments.

· Assisted researchers with grants by estimating sample size from power analysis using available expression data.

Bioinformatics Programmer - Gladstone Institutes — July 2018–July 2019

· Performed integrated multi-omic analyses, encompassing expression analysis, ATAC-seq, ChIP-seq, and motif enrichment, to uncover regulatory mechanisms in biological data for 20+ labs.

· Applied GLM, SVM, random forest and other advanced statistical and machine learning methods for deeper insights.

Education

Ph.D. Bioinformatics - University of California, San Francisco

· Statistics and machine learning to identify enhancers and genes predictive of tissue developmental phenotypes.

· Predicted enhancers in limb development integrating epigenetic data and modeled transcription factor binding.

B.S. Biochemistry - University of California, Los Angeles

Publications

Cordero, G.A., Holloway, A.K., Friedrich, T., Eme, J., Eckalbar, W., Kusumi, K., Janzen, F.J., Hicks, J.W., Conlon, F.L., Bruneau, B.G., Pollard, K.S. **Comparative transcriptomics of the heart illustrates challenges to extending classical theory to gene expression trends in animal development.** Genes, Development and Evolution (Manuscript in preparation).

Friedrich, T.*, Betty, B.*, Mason, M., VanderMeer, J., Zhao, J., Eckalbar, W., Logan, M., Pollard, K. S., Illini, N., Ahituv, N. **Bat accelerated regions identify a bat forelimb specific enhancer in the HoxD locus**. PLoS Genetics 12(3):e1005738 (2016). * Co-first authors.

Eckalbar, W., Schlebusch, S., Mason, M., Gill, Z., Booker, B., Nishizaki, S., Nday, C., Terhune, E., Nevonen, K., Makki, N., Friedrich, T., VanderMeer, J., Pollard, K.S., Carbone, L., Wall, J., Illing, N., Parker, A., Ahituv, N. **Genomic characterization of the developing bat wing.** Nature Genetics 48, 528-536 (2016).

Myers, S.A., Petted, S., Chatterjee, N., Friedrich, T., Tomoda, K., Krings, G., Thomas, S., Broeker, M., Maynard, J., Thomson, M., Pollard, K., Yamanaka, S., Burlingame, A. L., Panning, B. **SOX2 O-GlcNAcylation alters its protein-protein interactions and genomic occupancy to modulate gene expression in pluripotent cells.** eLife 5:e10647 (2016).

Kostka, D., Friedrich, T., Holloway, A. K., Pollard, K.S. **motifDiverge: a model for assessing the statistical significance of gene regulatory motif divergence between two DNA sequences.** Statistics and Its interface 8(4), 463–476 (2015).

Fogel, B. L., Wexler, E., Wahnich A., Friedrich T., Vijayendran C., Gao F., Parikshak N., Konopka G., Geschwind D.H. **RBFOX1 Regulates Both Splicing and Transcriptional Networks in Human Neuronal Development.** Hum. Mol. Genet. 1–16 (2012).

Konopka, G., Friedrich, T., Davis-Turak, J., Winden, K., Oldham M.C., Gao, F., Chen, L, Wang, G., Luo, R., Preuss, T.M., Geschwind, D.H. **Human-Specific Transcriptional Networks in the Brain.** Neuron 75, 601–617 (2012).

Schein, S., Sands-Kidner, M., Friedrich, T. **The physical basis for the head-to-tail rule that excludes most fullerene cages from self-assembly.** Biophys. J. 94, 938–57 (2008).

Schein, S., Friedrich, T. **A geometric constraint, the head-to-tail exclusion rule, may be the basis for the isolated-pentagon rule in fullerenes with more than 60 vertices.** Proc. Natl. Acad. Sci. U. S. A. 105, 19142–7 (2008).

Schein, S., Friedrich, T. **A geometric constraint, the head-to-tail exclusion rule, may be the basis for the isolated-pentagon rule in fullerenes with more than 60 vertices.** Proc. Natl. Acad. Sci. U. S. A. 105, 19142–7 (2008).

Nagarajan, S., Friedrich, T., Garcia, M., Kambham, N., Sarwal, M.M. **Gastrointestinal leukocytoclastic vasculitis: an adverse effect of sirolimus.** Pediatr. Transplant. 9, 97–100 (2005).