Yue (Jason) Zhao

www.linkedin.com/in/jsnzhao
yuezh@bu.edu
857-800-5610

Education

Boston University, Boston, MA *Sept. 2016 - Present*
PhD student in Bioinformatics, School of Engineering, BUSM

Shanghai Jiao Tong University, Shanghai, China *Sept. 2011 - Jun. 2015*
B.S. in Plant Biotechnology

University of California, Los Angeles, Los Angeles, CA *Jul. 2014 - Sept. 2014*
Fellow of Cross-disciplinary Scholar in Science and Technology (*CSST*) Program

Experience

Boston University School of Medicine, Boston, MA
Research Assistant in Section of Computational Biomedicine *Sept. 2016 - Present*
- Performed feature engineering by differential expression analysis and elastic net for Tuberculosis(TB) progression
- Built random forest and SVM model for TB progression classification based on selected genes from elastic net

Duke University, Durham, NC
Research Assistant in Department of Electronic and Computer Engineering *Sept. 2015 - Mar. 2016*
- Modified Poisson Factor Analysis model to dynamic with Bernoulli-Poisson Link
- Applied Dynamic Poisson Factor Analysis model to extract latent structure from microbiome data

Publications

[1] Y. Zhang, **Y. Zhao**, L. David, R. Henao, L. Carin "**Dynamic Poisson Factor Analysis**", in *IEEE International Conference on Data Mining* (2016), doi: 10.1109/ICDM.2016.0186

[2] C. Sun. et al "**RPAN: rice pan-genome browser for 3000 rice genomes**", in *Nucleic Acids Research* (2016), doi: 10.1093/nar/gkw958

[3] Z. Hu. et al. "**EUPAN enables pan-genome studies of a large number of eukaryotic genomes**", Accepted by *Bioinformatics.*

Projects

Modeling Tuberculosis (TB) Progression with Baseline Transcriptomic Signatures *Sept. 2016 - Present*
- Implemented data cleaning, quality control, batch visulization and correction on multiple RNA-seq datasets
- Identified a set of 42-gene signature achieving 0.92 AUC in isolated testing set classification prediction

Microbiome Profiling in the Airways of Lung Cancer Patients *Feb. 2017 - Present*
- Implemented Pathoscope to quantifying the proportions of microbiome reads from total airway RNA-seq data
- Developing nasal genomic classifier for the diagnostics of lung cancer combining microbiome and host transcriptome

Gene Signatures of Immune Responsiveness to Brain Cancer Chemo-immunotherapy *Nov. 2016 - Present*
- Developed a pipeline for genomic feature selection combining differential expression and random forest
- Identified a 6-gene signature separating responsive and unresponsive glioma using lasso logistic regression

Topic Modeling in Metagenomics Data *Sept. 2015 - Mar. 2016*
- Modified the dynamic PFA model with stochastic MCMC methods and GPU parallel computing using CUDA
- Identified topics of OTUs from human metagenomic data with potential biological function enrichment

3,000 Rice Pan-genome Construction *Sept. 2014 - Aug. 2015*
- Designed pipeline for unaligned sequence prediction and RNA-seq verification of discovered genes.
- Built local database containing pan-genome data and designed user interface for website using Mysql

Programming Languages

Proficient in: R, Python, Shell, LaTeX, MATLAB, C/C++
Exposure to: CUDA, PHP, Perl, MySQL, HTML, CSS