Li Shen

Data Scientist, Bioinformatician

An oncology-focused data scientist with a 15+ year proven track record of developing and productionizing data products to advance treatment and diagnosis.

Contact

Email: lshen7@asu.edu
Ph: 480-452-5464
Loc: Houston, TX
LinkedIn
Web Portfolio
ORCiD

Education

PMS - Computational BioScience, Arizona State
MS - Engineering, Arizona State

Skills

R & Shiny (Tidyverse)
Python (Numpy, Pandas, Scikit-Learn)
Unix/Linux HPC environment
AWS & Cloud
Snakemake
Git & Docker/Singularity
ML (Clust., Reg., Class., Time Series.)
AI (CNN, RNN)

Work Experience

Senior Bioinformatician, Bioinformatics & Computational Biology

The University of Texas MD Anderson Cancer Center
February 2020 - Present

- Led the clinical development of robust pipelines for processing diverse in-house datasets obtained from multi-omics data platforms, encompassing bulk-RNA-seq, DNA-seq (WGS; WES), scRNA-seq (Single Cell), ChIP-seq, ATAC-seq, spatial transcriptomics and RPPA (Protein), among others, to investigate mouse, human and PDX models.
- Pioneered a systematic strategy to optimize in-house Drug-Screening processes, including the handling of single and combination drugs, as well as conducting comprehensive quality control and downstream analyses to identify druggable targets.
- Conducted complex analyses on in-house and large-scaled clinical and molecular datasets, leveraging resources such as TCGA, CCLE, GDSC, GEO and cBioPortal, resulting in the identification of critical biomarkers and the execution of integrative analyses utilizing a range of statistical models such as supervised/unsupervised clustering, regression, classification, and survival analysis.
- Proficiently annotated and interpreted results employed diverse databases and bioinformatics tools, including IGV, COSMIC, GO, KEGG, GSEA/ssGSEA and Cibersort deconvolution, enhancing the precision of data insights.
- Providing computational support to researchers and clinicians using diverse bioinformatics tools and methods. Independently managing complex tasks within deadlines and maintaining research computing tools.
- Demonstrated expertise in experimental design and the calculation of sample size and statistical power.
- Provided invaluable support for numerous multi-million-dollar grants at various stages, including RO1, SPORE, Moonshots and CPRIT, through grant application and ongoing data analyses.
- Co-authored more than 40 peer-reviewed journals, with over 90% of them having an impact factor exceeding 5 and more than 50% exceeding 10.
- Sample Project 1: SCLC Subtypes - Cancer Cell 2021 (R, Shiny)
- Sample Project 2: Single Cell Analysis Viewer (R, Shiny)

Professional Appointments

The University of Texas MD Anderson Cancer Center

- February 2020 - Present **Senior Bioinformatician**
- October 2014 - 2020 **Senior Statistical Analyst**
- June 2011 - 2014 **Research Statistical Analyst**
- April 2008 - 2011 **Statistical Analyst**