Lois Randolph, M.S.

lois.randolph@outlook.com | www.linkedin.com/in/lois-randolph | www.github.com/Low-is | https://low-is.github.io/Lois_portfolio/

SUMMARY

Bioinformatician with 3+ years of experience in multicohort biomarker discovery and predictive modeling from multi-omic and clinical data. Strong expertise in meta-analysis, advanced statistical modeling, and machine learning, with a focus on reproducible, scalable pipelines and cross-functional collaboration to translate complex analyses into actionable insights. ORCiD: https://orcid.org/0000-0001-8292-2639

SKILLS

Programming Languages: R, Python, SQL (HiveQL), SAS
Machine Learning: Scikit-Learn, Random Forest, Gradient Boosting, SVM, kNN
Deep Learning: TensorFlow, Keras, PyTorch
Data Engineering & Databases: Hadoop, MySQL, RSQLite
Workflow & Reproducibility: Docker, NextFlow, Git. Bash (Unix/Linux)
Visualization & Reporting: Power BI, Python, R
Development: Jupyter, RStudio, VS Code, PyCharm

WORK EXPERIENCE

UT Health Science Center   **Dec. 2023 - Present**
Bioinformatician   *San Antonio, TX*
- Responsible for leading the development of early diagnostic and prognostic models for neonatal and pediatric care to identify risk of disease onset and clinical outcomes.
- Integrated heterogeneous datasets including EMR, bulk RNA-seq, single-cell RNA-seq, spatial transcriptomics, methylation, microbiome, proteomics, and other omics or clinical data.
- Developed and applied statistical models ranging from simple linear regressions to mixed-effects frameworks on both static and longitudinal datasets, rigorously controlling for covariates and batch effects.
- Reproducible, containerized workflows that leverage parallel processing, optimizing data handling, harmonization, multicohort meta-analyses, and regression and classification tasks.
- Unsupervised and supervised machine learning,
- Refined single-cell type annotations and integrated single-cell with spatial transcriptomics mapping predicted cell types to lung spatial landscape.
- Version control with Git.
- Review potential grants for methodological rigor and data integrity; conducted exploratory analyses to identify errors and gaps prior to submission.

UT Health Science Center   **Aug. 2021 – Dec. 2023**
Graduate Research Assistant   *San Antonio, TX*
- Processed raw bulk RNA-seq FASTQO files; quality assessment, adapter and low-quality base trimming, alignment to mouse (mm9) and human reference genomes, and generation of gene-level expression count matrices.
- Processed raw single-cell RNA-seq FASTQ files to gene-by-cell expression matrices; demultiplexing, barcode and UMI processing, quality control, read alignment to reference genomes or transcriptomes, and feature count matrix generation.

- Differential expression analysis on triple-negative breast cancer (TNBC) xenograft and syngeneic mice models, uncovering suppression of oncogenic signaling pathways in high-fat diet (HFD) groups treated with the small-molecule compound EC359.

Baylor College of Medicine   June 2021 – Aug. 2021
Intern   *Houston, TX*
- Enhanced data curation and pipeline efficiency for gene and variant information, supporting large-scale research initiatives.
- Translated Ruby code to JavaScript to improve API interactions and data parsing across platforms.

UT Health Science Center   June 2020 – June 2021
PREP Scholar   *San Antonio, TX*
- Automated image analysis workflows for protein expression studies, improving annotation accuracy and throughput.
- Applied statistical and image processing techniques to quantify protein interactions and expression patterns.
- Image processing in Python to enhance microscopy images, improving quality and contrast, annotating labels, and tailoring visuals for professional reporting.

PROJECTS

Computational Biomarker Discovery in Neonatal and Pediatric Disease

- Integrative analyses across spatial, GWAS, and clinical datasets, investigating disease mechanisms and identifying candidate biomarkers in neonatal and pediatric populations.

Multi-image Classification with 3D CT scans Using CNN

- Trained 3D convolutional neural network (CNN) to classify lung cancer histological subtypes from volumetric CT scan data, demonstrating applied deep learning on 3D medical imaging.
- Achieved 75% training accuracy and conducted validation analysis, reporting 31% validation accuracy and associated loss metrics to assess generalization for future improvements.

EMR Data Engineering Pipeline

- Scalable ETL pipelines for healthcare data, leveraging Python, Hive, and Hadoop.
- Extract, transform, and load complex EMR datasets into a data warehouse environment for analytics and reporting.

EDUCATION

M.S. in Cancer Biology   May 2023
University of Texas Health Science Center San Antonio   *San Antonio, TX*

B.S. in Biology | Minor in Mathematics   May 2020
University of Texas at San Antonio   *San Antonio, TX*