Shalini Karthyk

• +1 617-373-0081 • karthyk.shalini@gmail.com • LinkedIn • Portfolio • Boston, MA

SUMMARY

Experienced bioinformatics and data science professional with nearly **5 years** of diverse expertise spanning academia and industry. Proven in data engineering, bioinformatics pipelines, and real-world data analytics. Skilled in cross-functional collaborations, translating business needs into actionable requirements, and driving data-driven innovative solutions in biotech and pharma R&D.

EDUCATION

Master of Science in Bioinformatics | Graduate Certificate in Data Analytics **January 2022 - June 2024**
Northeastern University, Boston, MA **GPA: 4.0/4.0**

Bachelor of Technology in Bioinformatics **July 2015 - July 2019**
SASTRA University, Tanjore, TN, India

TECHNICAL SKILLS

Bioinformatics Databases: BLAST, NCBI, GenBank, PDB, EMBL, Entrez, SwissProt, UniProt, ENCODE, GENCODE, Ensembl.
Genomic Analysis: GWAS, Exome, Whole Genome, Transcriptomics, Epigenetics, Structural Variants, Functional Genomics.
Programming Languages: R, Python, SQL, NoSQL, MATLAB, C, C++, Java, Perl, PHP, HTML, Shell script.
Machine Learning: Classification (DT, Random Forest, SVM, KNN, Neural Networks), Regression (Ridge, Lasso), Clustering.
Computing: High-Performance Computing (Linux/Unix, Cluster), Cloud (AWS, Code Ocean), Docker, Version Control(Git/GitHub).
Project Management: Trello/JIRA, Benchling, LabArchive

PROFESSIONAL WORK EXPERIENCE

The Institute for Experiential AI **Boston, MA**
Bioinformatics and Machine Learning Analyst Co-op **January 2024 – Present**
- Developed **ETL pipeline** for ENCODE shRNA RNA-binding proteins (RBP) Knockdown data, improving differential splicing analysis by 60% for over 10,000 splice junctions.
- Orchestrated Machine Learning pipeline integrating **XGBoost** and **SHAP** for robust biological data interpretation, enhancing predictive accuracy by 20%.
- Engineered **novel graph structure** for analyzing mRNA splicing regulation by RBPs, transforming post-transcriptional gene expression control understanding for 200+ RBPs and advancing genomics research.

Northeastern University **Boston, MA**
Data Validation Assistant **January 2022 – December 2023**
- Streamlined output tool for ALERT CLASP video analytics project research data with a revamped data structure using **MATLAB** and advanced algorithms, reducing report generation time by **50%** and enhancing usability.
- Automated video annotation software using **advanced tracking algorithms**, resulting in a **70%** increase in productivity.
- Engineered robust QA measures, enabling re-coding of 50+ bugs and a 30% decrease in overall bug recurrence rate.

Fulcrum Therapeutics **Cambridge, MA**
Bioinformatics Associate Intern **June 2023 – August 2023**
- Pioneered a **high-throughput** erythroid differentiation culture system, increasing CRISPR and small molecule screening for drug development efficiency by 30% across 150 compounds accelerating **hematological disease investigation**.
- Devised proprietary **R Shiny** web-based Interface algorithm to identify erythroid lineage markers using single-cell and qPCR, with 40% higher precision than existing methods, reducing potential targets to top 3 sure marker genes.
- Optimized flow cytometry for cell co-positivity analysis and sorting, implementing a novel gating strategy, enhancing 35% cell population purity.

Brigham and Women's Hospital **Boston, MA**
Technical Research Assistant **April 2021 - September 2021**
- Identified 8 nephrotoxicity-related genes (DMT1, GLUT1, etc.) in kidney organoids through comprehensive genomic and epigenetic analysis of 20 million sequencing reads using R.
- Implemented **NGS data** analysis data pipelines to examine kidney organoid transcriptomes, revealing 1,500+ genetic variations and profiling gene expression levels for 7,000 genes.
- Improved **multiplex assay** efficiency by 25% for 18 biomarkers using Luminex, MSD, and Randox, retaining 90% sensitivity.

Research Trainee **September 2019 - September 2020**
- Led a team of 3 in devising the **CRISPR genome-editing** tool to excise 12 exons in Duchenne Muscular Dystrophy gene.
- Enhanced **SARS-CoV-2** variant detection by 40% through implementing **immuno-PCR** techniques, and isolating immune cells and specific antibodies against COVID-19.
- Discovered 3-fold Rho-Rac activity increase in EDA fibronectin pathway through GST pull-downs of 10 Rho proteins.

Harvard Medical School, Department of Genetics **Boston, MA**
Research Intern **February 2019 – June 2019**
- Established an **R-Seurat pipeline** for single-cell RNA Seq kidney organoid data and mapped cell fate decisions using Monocle.
- Achieved CRISPR/Cas9 mediated mutation of 6 kidney disease genes in iPSC lines and rescued disease phenotype, quantifying **post-treatment gene expression** changes across 5000 genes by Bulk RNA-seq analysis.
- Dissected molecular circuits governing nephron progenitor cell development with scalable RNA profiling using Perturb-seq.
- Enhanced RNA-seq alignment and genetic screen analysis with **BLAST, BWA, and samtools**, improving data interpretation.