

PRANAVA UPPARLAPALLI

pxu.bioinfo@gmail.com | LinkedIn | GitHub | +1 (945) 527-5490 | Leesburg, VA 20175

Summary

Bioinformatics scientist (MS) who transforms large-scale biological data into actionable insights by building scalable, data-agnostic pipelines with Python, R, and Nextflow. I am seeking a role where I can apply my computational expertise to accelerate research and development in a dynamic industry environment.

Education

- University of Texas at Dallas, Richardson, TX** Aug 2023 – May 2025
Master of Science in Bioinformatics and Computational Biology
- Bharati Vidyapeeth University, Pune, MH, India** Aug 2022 – May 2023
Advanced Diploma, Bioinformatics
- Sri Venkateshwara University, Tirupati, AP, India** June 2019 – May 2022
Bachelor of Science in Microbiology, Biochemistry, and Chemistry

Work Experience

- Graduate Student Researcher, Xuan lab, The University of Texas at Dallas** Jan 2025 – May 2025
 - Tested a GWAS-based machine learning algorithm based on PrediXcan using genomic data to accurately predict gene expression in different tissues, leveraging eQTL models and GTEx database.
 - Analyzed the role of trans SNPs using HI-C data, to decipher the relationship between genotype and phenotype using SNP variants.
 - Created pipelines and provided efficient data analysis using Bash scripting and HPC computing.
- Undergraduate Research - Sree Vidyannikethan Degree College** Aug 2020 – Mar 2021
 - Researched and characterized the antioxidant properties of *Biancaea sappan*, employing advanced biochemical assays for compound analysis.
 - Designed and executed lab experiments that increased compound yield efficiency by 15%.
 - Optimized growth conditions by formulating specialized media and conducting antibacterial assays.

Projects

- Cancer RNA-Seq Expression Analysis** May 2025 – Jun 2025
 - Explored gene expression across five major cancer types using RNA-Seq data. Built a complete R-based pipeline to perform QC, normalization, and differential expression analysis using DESeq2.
 - Generated heatmaps, PCA plots, and volcano/MA plots to identify biomarkers and clustering patterns across BRCA, KIRC, LUAD, COAD, and PRAD.
- SeqMorph: Sequence Mutation Simulator** Sep 2024 – Jan 2025
 - Engineered a versatile bioinformatics tool in Python to simulate diverse mutations (substitutions, insertions, deletions) in DNA, RNA, and protein sequences.
 - Implemented features for mutation tracking, sequence analysis (e.g., codon changes, frameshift detection), and exporting results in FASTA/CSV formats for downstream use.
- Yeast-Stress Analysis** Dec 2024 – Jan 2025
 - Designed a reproducible RNA-Seq analysis pipeline using Nextflow and Docker, incorporating Fastp for QC, HISAT2 for alignment, SAMtools for BAM processing, and FeatureCounts for quantification.
 - Performed differential expression analysis using edgeR to investigate transcriptomic changes in yeast under oxidative stress, followed by pathway enrichment and visualization with volcano plots and heatmaps.
- Gleason Score Classification Using ResNet-50** Sep 2024 – Dec 2024
 - Designed a deep learning pipeline using ResNet-50 for prostate cancer classification, achieving 90% classification accuracy.
 - Implemented data augmentation techniques to improve model robustness and generalizability.

Skills

Programming & Databases	Python, R, Bash/Shell Scripting, SQL
Machine Learning	Scikit-learn, PyTorch, Pandas, NumPy, Deep Learning (CNNs, ResNet)
Workflow & Cloud Computing	Nextflow, Docker, Git, Conda, Slurm, AWS (S3, EC2)
Sequencing & Genomics	RNA-Seq, scRNA-Seq (Seurat, Scanpy), ChIP-Seq, Variant Calling (GATK), Samtools, FastQC, BEDTools
Genetic & Statistical Analysis	GWAS, TWAS, eQTL Analysis, Gene Set Enrichment (GSEA), PLINK
Bioinformatics Resources	NCBI, IGV, Geneious, Ensembl, UCSC Genome Browser, GEO, dbSNP, SRA, KEGG

Relevant Courses

- Applied Bioinformatics, Analysis of Data Structures and Algorithms, Big Data, Molecular Biology, Biochemistry, Probability & Statistics