PANKAJ PRADEEP

65 West 106th Street
New York, NY 10025
https://www.linkedin.com/in/pankajpradeep/
https://github.com/PankajPradeep/BRCA-SAE
(332) 250-4539
pp2831@columbia.edu

EDUCATION

Columbia University — New York, NY
Master of Science in Biomedical Engineering, 3.57/4 — Expected Dec 2023
Coursework: Machine Learning for Functional Genomics, Statistical Machine Learning for Genomics, Computational Modeling of Physiological Systems, Biomedical Innovation, Applied Data Science

Vellore Institute of Technology (VIT) — Vellore, India
Bachelor of Technology in Biotechnology, 9.05/10 — July 2022
Coursework: Bioinformatics (Genetic Alignment and Protein Engineering), Molecular Biology, Immunology and Immunotechnology, Genetic Engineering, Programming (Python, R, MATLAB)

SKILLS

Programming	**R/RStudio, UNIX, Python,** MATLAB, **Git, Machine Learning,** Bash Scripting, **HPC**
Bioinformatics	Seurat, Bioconductor, scRNAseq, bulk RNAseq, Perturbed Cancer Data, **Omics**
Protein Visualization	Mutant, Polyphen, PyMol, SwissPDBViewer, SIFT
Databases	GEO, TCGA, CCLE, UCSC, PubMed, UniProt, NCBI, EMBL-EBI
Other	Microsoft Office, Google Docs, **LaTeX**, Mendeley, Photoshop
Lab Skills	Genomic Data Analysis, Plant Tissue Culture, DNA/RNA Extraction, PCR, Microbiology
Packages	Scikit-learn, Pandas, NumPy, Keras, Tensorflow, Seaborn

EXPERIENCE

Gertrude H. Sergivesky Centre, Columbia Irving Medical Centre — New York, NY
Student Research Worker *under Dr. Giuseppe Tosto* — Sep 2023 - Present
- Pre-processing FASTQ files for Bioinformatics Analysis for studying Alzheimer Disease
- Evaluating AD risk factors in underrepresented populations using R and Unix
- Investigating whole genome sequenced and GWAS data for differences in populations based on their ancestry (genetic composition) and risks for each of those groups
- Performing deconvolution of transcriptomics data from RNAseq to single cells

GlaxoSmithKline Pharmaceuticals (Atrium Payroll Services Inc.) — Collegeville, PA
Summer Intern – Computational Biology Oncology — Jun 2023 – Aug 2023
- Investigated clinical signatures for cancer gene/protein expression from public databases
- Systematically evaluated correlation patterns between RNA expression and proteomics in cancer cell line data from Cancer Cell Line Encyclopedia and TCGA using Python
- Ingested and processed scRNAseq cancer datasets from GEO into Bioinformatics pipelines

Herbert and Florence Irving Institute for Cancer Dynamics, Columbia University — New York, NY
Student Research Worker — Oct 2022 – June 2023
- Conducted Genomic Data Analysis to identify role of extrachromosomal DNA in the evolution of cancer, using UNIX, Linux and RStudio
- Analyzed 50 samples for extrachromosomal DNA using publicly available tools on GitHub
- Assisted in setting up Packages for analyzing gene data of esophageal cancer
- Discovered novel regions of DNA that could potentially be ecDNA hits while collaborating with Professor Karol Nowicki-Osuch and a PhD student

RESEARCH PROJECTS

Columbia University: Predictive Model for Alzheimer's Disease Prediction using Machine Learning — New York, NY
- Collected RNA-seq data for Alzheimer patients from publicly available databases — Jan 2023 – May 2023
- Compared Normal vs Diseased State Gene expression using ANOVA test
- Extracted top Principal Components and Stratified Dataset to get training and test datasets
- Explored various Predictive Models and chose one that had least false negatives

Columbia University: Multi-Omics Data Integration using Stacked Autoencoders — New York, NY
- Extracted and Processed RNAseq, Methylation and Proteomics Data for breast invasive carcinoma — Jan 2023 – May 2023
- Built a Stacked Autoencoder model that represents multi-omic integrated data in a latent space
- Performed Feature Attribution to discover biomarkers that are drivers of cancer progression across all datasets
- Evaluated top 20 biomarkers and their functions by performing extensive literature survey to confirm our results

INTERESTS

Tennis, Running, Soccer, Hack-a-thons, Cancer Research, Medical Devices, Point-of-care diagnostics, Cancer Therapy