Ketaki Prakash Ghatole

📞 412-892-1861 ✉ kghatole@andrew.cmu.edu 🔗 www.linkedin.com/in/ketaki-g 🪪 https://www.ketaki-ghatole.com/

Education

Carnegie Mellon University **Aug 2021 – May 2023**

Master of Science in Computational Biology; GPA: 3.92/4
Relevant Courses: Mathematics and Statistics, Machine Learning, Neural Network and Deep Learning, Quantitative and Population Genetics, Computational Molecular Biology, Computational Medicine, Computational Genomics

Ramaiah Institute of Technology **Aug 2016 – Aug 2020**

Bachelor of Engineering in Biotechnology; GPA: 9.25/10; Best Project Award 2020
Relevant Courses: Cell Biology, Microbiology, Structural Biology, Genomics, Proteomics, Bioinformatics, Biostatistics

Experience

Moderna Services Inc. **Jun 2022 – Aug 2022**

Computational Science Intern *Cambridge, MA*
- Implemented a bioinformatics pipeline using DolphinNext to process e-CLIP data from ENCORE and obtain the peaks. Analyzed and visualized this data in Python
- Developed a web-based application using Python Shiny to predict and visualize potential binding motifs for RNA binding proteins

Schwartz Lab, Carnegie Mellon University **Jan 2022 – May 2022**

Graduate Research Assistant *Pittsburgh, PA*
- Analyzed RNA-seq data from patients suffering from hematological malignancies under-going CAR-T cell therapy
- Developed a machine learning framework to the identify key features predictive of relapse

Axiom Healthcare Strategies **Oct 2020 – Mar 2021**

Junior Analyst *Bengaluru, India*
- Analyzed data from clinical trials, press releases, and conferences to predict the competitive landscape in the Oncology
- Automated data collection and data cleaning for monitoring projects using web scraping techniques in Python to improve the work efficiency

Projects

Malaria Detection using ResNet50 Architecture **Feb 2023 - Mar 2023**
- Developed a deep learning model with 97% test accuracy using transfer learning with the ResNet50 architecture to detect malaria parasites in blood smear

Multi-omic Analysis of Lung Cancer using Autoencoders **Sep 2022 - Dec 2022**
- Implemented auto-encoders using Pytorch and for unsupervised feature selection of gene expression and CNV data.
- Classified Lung Cancer using SVM and Random Forest using mono-omic and multi-omic data with an accuracy of 95%

Prediction of RNA-protein interactions in SARS-CoV2 🐙 **Mar 2022 - May 2022**
- Trained a neural network-based classifier to predict the interactions between host cell proteins and viral RNA sequences from SARS-CoV-2 using e-CLIP data from ENCORE

Evaluation of machine learning models for classification of Glioma 🐙 **Jan 2022 – May 2022**
- Built four supervised multiclass classifiers (SVM, LR, KNN, Naive Bayes) to identify the type of glioma given gene expression data from TCGA
- Compared their efficiency using 5-cross validation and obtained the highest accuracy of 96% with SVM

Computational Analyses to determine Codon Usage Bias 🐙 **Nov 2019 – Aug 2020**
- Analyzed the amino acid preferences and codon usage bias using sequencing data from VenomZone, Expasy
- Identified differences in the evolution of genes coding for venom proteins and the genes coding for non-venom proteins

Technical Skills

Languages and Framework: Python, R, Bash, SQL, Pandas, Biopython, Scikit-learn, Pytorch, Tableau
Bioinformatics: AutoDock, GATK, BLAST, Bowtie, BWA, Samtools, IGV, Bedtools, STAR
Tools and Technologies: Confluence, HPC, AWS, LaTex, Docker, Git, Visual Studio

Leadership and Awards

- James R. Swartz Entrepreneurial Fellow 2021 cohort at Carnegie Mellon University
- Co-founder and Graphic Designer for the logic and reasoning club "Numera" at Ramaiah Institute of Technology
- Volunteered at the NGO Janaseva Samithi to spread awareness about women and child literacy in rural India