Pratik Katte

Cell: +1 (831)-295-0420

pratikkatte7@gmail.in, pkatte@ucsc.edu

[Portfolio] [LinkedIn] [Github]

WORK EXPERIENCE

Graduate Student Researcher, Corbett-Detig Lab, Santa Cruz, CA   *Oct 2023 – Present*

- Currently exploring the application of Generative Flow Networks (GFNs) in improving phylogenetic tree assembly, focusing on their efficacy with large and intricate genomic datasets.
- Also involved in designing and developing a dashboard that helps in studying distinct virus lineages identified from SARS-CoV-2 RNA fragments extracted from wastewater samples. The dashboard leverages UShER from global-phylogenetic inference and Freyja for performing lineage deconvolution.

Senior Research Engineer, Niramai Health Analytix, Bangalore, India   *Jan 2020 – Jun 2023*

- Led the development of machine learning algorithms within a desktop application aimed at assisting technicians in the efficient capture of breast thermal images for cancer screening purposes. This application is currently deployed in over 100 hospitals worldwide.
- Played a lead role in developing an AI-based desktop application designed to screen individuals for COVID-19 symptoms, including elevated temperature and shortness of breath, while enforcing adherence to COVID-19 protocols. Successfully screened over 1 million individuals across India.
- Contributed to training a U-Net based deep learning model to perform lung segmentation on compressed chest X-ray images within the Xray-Setu project. Notably, this effort resulted in a substantial 6% enhancement in the AUC score of the lung abnormality classifier.
- <u>Awards</u>: "On Spot Award" and "Working beyond boundaries Award" for the performance on Niramai Fever Test product.

Co-founder, Rewind, Mumbai, India   *Sep 2018 – Nov 2019*

- Co-founded a venture Rewind that capitalized on established waste pickers, paper and metal shops, and recyclers network to create a platform enabling the sale of diverse scraps, including electronics and furniture.
- Operated in stealth mode for one year, collaborating with 10 waste pickers and 2 recyclers to collect electronics and furniture totaling $1000 in value.

Research Intern, L.V. Prasad Eye Institute, Hyderabad, India   *Jun 2018 – Aug 2018*

- Designed and developed a conversational chatbot using a decision tree algorithm to streamline booking an appointment with a doctor and educate patients about eye diseases. Deployed on a website serving more than 4500 visitors per day.
- Developed an algorithm to quantify the health of an eye using fuzzy c-means clustering algorithm.

Data Science Intern, Prakshep, Bangalore, India   *Jun 2017 – Jan 2018*

- Explored unsupervised machine learning techniques such as Gaussian Mixture Model, along with employing deep learning models like U-Net, for the segmentation of forested and deforested areas within satellite imagery..
- Created a Geographic Information System (GIS) using GeoServer and the GDAL Python library to process geospatial data pertaining to crop harvests. Empowered farmers to make informed decisions regarding seed selection for improved yields.

EDUCATION

University of California Santa Cruz, California, USA   *Sep 2023*

Masters in Biomolecular Engineering and Bioinformatics

Relevant Coursework: Applied RNA Bioinformatics, Bioinformatics - Models and Algorithms

University of Mumbai, Mumbai, India   *Jun 2015*

Bachelor in Engineering, Information Technology

Relevant Coursework: Intelligent System, Image Processing, Computer Graphics and Virtual Reality, Big Data Analytics

PUBLICATIONS

Pratik K., et al. 'Automated Thermal Screening for COVID-19 using Machine Learning.' arXiv preprint arXiv:2203.14128 (2022).

Sabyasachi S., **Pratik K.,** et. al. 'Abstract PS2-44: Diagnosing COVID-19 From Images of Chest X-rays Communicated Via WhatsApp.' UKIO Congress (2022).

Patil, Vivek, **Pratik Katte**, and Abhay Patil. 'Restoration of Images Using Only Noisy Data.' International Journal of Research and Analytical Reviews (IJRAR) 6.1 (2019).

CONFERENCE TALKS

- 'Niramai Fever Test: Automated Screening for COVID Symptoms', Wolfram Technology Conference – 2021 [link]
- 'Machine learning for COVID-19 detection', Data Science Conference, Europe – 2021 [link]

PROJECTS

Phylogenetic Tree Reconstruction
- I am actively conducting research on various computational methods for phylogenetic tree reconstruction, focusing on assessing their effectiveness in handling densely populated phylogenetic trees.
- And also working on the integration of Generative Flow Networks (GFNs) into phylogenetic tree reconstruction and assessing its feasibility in handling massive, dense genomic datasets.

Gene Expression Analysis of Small RNA
- Vibrio cholerae is responsible for cholera, the role of small RNAs(sRNAs) remains inadequately understood. Investigating the role of small RNA molecules in communicating with other colonies holds promise for gaining a deeper understanding of the root causes of infectious diseases caused by human pathogens.
- Using the OTTR seq dataset of Vibrio Cholerae and tRAX software tool, we perform gene expression analysis between the stationary and biofilm state.

StructHunt [github]
- As a part of a hackathon at UCSF, we created a tool designed to track the publication of new research papers detailing integrative biomolecular structures in bioRxiv and medRxiv. This tool enables us to swiftly capture and incorporate this valuable new data into the RCSB Protein Data Bank.
- I was responsible for the implementation of context retrieval from LLM embeddings, utilizing the FAISS library for efficient indexing and similarity searches.

TECHNICAL SKILLS

- Bioinformatics Tools: Sam Tools, UCSC Genome Browser.
- Programming Languages: Python, NodeJS, ReactJS, C++, C#.
- Frameworks/Libraries: TensorFlow, PyTorch, Keras, Django, WPF