PRAPHULLA BHAWSAR

AI Engineer / PhD Student

https://prafulb.github.io | (929) 409-9237

github.com/PrafulB

LinkedIn.com/in/PrafulB

praphulla.bhawsar@stonybrook.edu

Education

- **PH.D. in Bioinformatics** | ***Stony Brook University, NY*** — AUG 2021 - Present
- **MASTER OF SCIENCE** | ***New York University, NY*** — AUG 2017 - MAY 2019
  Major: Computer Science / Data Science

Experience

DOCTORAL FELLOW | ***National Cancer Institute, Rockville, MD*** — AUG 2021 - Present

DATA ENGINEER | ***National Cancer Institute, Rockville, MD*** — MAR 2019 – AUG 2021

- Working with the Data Science & Engineering group to design AI tools for the edge to enable FAIR cancer research.
- Created dashboards to track excess mortality due to COVID-19 using real-time data from multiple sources.
- Conducting research on applying AI to digital pathology and radiology imaging data to assist domain experts in their analyses.

LEAD DEVELOPER | ***NYU Student Tech Innovation, NY*** — SEP 2017 - MAR 2019

- Led the development team of the Student Tech Innovation team at NYU to create novel applications for the NYU community.

FULL STACK DEVELOPER | ***Reliance Jio Infocomm Ltd., India*** — JUN 2015 - JUL 2017

- Lead backend developer for the JioTV2.0 live TV streaming application, currently used by over 100 million people in India.
- Redesigned the backends of several apps as part of the core R&D team; also worked on multiple PoCs and exploratory projects.

Projects

- **ImageBox3: Client-side Whole Slide Tile Serving:** Developed a completely in-browser tiling mechanism to view and operate upon digital pathology whole slide images (WSI). Leveraged the similarities between WSI data formats and cloud-optimized GeoTIFF to allow for on-demand access to whole slide regions with no server deployment required.
- **epiPath: Digital Pathology in the Browser:** Designed an open-source, platform to orchestrate the entire medical image analysis pipeline in the browser, from raw data to annotation to model inference, entirely in the web browser. This zero-footprint application has no server-side, can use imaging data hosted on any cloud storage service and leverages TensorFlow.js to allow powerful neural networks to be run on the local machine at zero cost.
- **Galaxy Morphology Classification Using CNNs:** Trained a VGG-19 CNN using the Keras API to classify galaxies by morphology from raw sky survey images. Compared the performance of an uninitialized model to one that was trained on another dataset first, demonstrating the advantages of using Transfer Learning on sparse astronomical datasets.

Technical Skills

Languages & Frameworks : JavaScript, Node.js, Python, Golang, TypeScript, ReactJS, HTML, CSS, PHP

Databases & Misc: TensorFlow, TF.js, PyTorch, SQL, AWS, GCP, OpenCV, MongoDB, Elasticsearch, Redis, Couchbase, Firestore, BigQuery, Tableau, D3.js, Plotly.js, Nginx

Publications

- Almeida, J. S., Shiels, M., Bhawsar, P., Patel, B., Nemeth, E., Moffitt, R., Closas, M. G., Freedman, N., & Berrington, A. (2020). Mortality Tracker: the COVID-19 case for real time web APIs as epidemiology commons. Bioinformatics.
- Zhang, T., Joubert, P., Ansari-Pour, N., et al. Genomic and evolutionary classification of lung cancer in never smokers, Nature Genetics (2021).
- Bhawsar PM, Almeida JS et al. Browser-based data annotation, active learning and real-time distribution of artificial intelligence models: from tumor tissue microarrays to COVID-19 radiology, Journal of Pathology Informatics (2021).