PRAPHULLA BHAWSAR

github.com/PrafulB

LinkedIn.com/in/PrafulB

AI Engineer | PhD Student

https://prafulb.github.io | (929) 409-9237 praphulla.bhawsar@stonybrook.edu

Education

- **Ph.D in Biomedical Data Science** | ***Stony Brook University, NY*** AUG 2021 - Present
- **M.S. in Computer Science** | ***New York University, NY*** AUG 2017 - MAY 2019

Experience

PREDOCTORAL 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 at the Division of Cancer Epidemiology and Genetics to design Findable, Accessible, Interoperable and Reusable AI tools for 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. Demonstrated to support 100 million+ concurrent users.
- 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 formats and cloud-optimized GeoTIFF data to allow for on-demand access to slide regions without a server.
- **epiPath: Digital Pathology in the Browser:** Designed a zero-cost, zero-footprint platform to orchestrate the complete medical image analysis pipeline, from annotations on raw data to visualizing model predictions to iterative active learning, entirely in the user's web browser under their own governance.

Technical Skills

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

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

Publications

- Almeida, J. S., Shiels, M., Bhawsar, P., et al. Mortality Tracker: the COVID-19 case for real time web APIs as epidemiology commons. Bioinformatics (2020)
- 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)

- Almeida JS, Patel B, et al. epiDonate – distributed serverless data infrastructure for epidemiological studies, AMIA Jt. Summits Transl Sci Proc. (2023)
- Bhawsar PM, Bremer E et al. ImageBox3: No-Server Tile Serving to Traverse Whole Slide Images on the Web, Under peer review (2023)
- Sandoval L, Jafri S, et al. PRScalc, a privacy-preserving calculation of raw polygenic risk scores from direct-to-consumer genomics data, Bioinformatics Advances (2023)
- Martim Afonso, Bhawsar PM et al. Finding Regions of Interest in Whole Slide Images Using Multiple Instance Learning, arXiv (2024)