Budhachandra Singh Yumkhaibam

bxy4155 | YBCS | Website | budhachandra

Research Interest

Medical Imaging . Machine Learning for Mental Health · Multi-Modal Integration . Interpretable AI in Healthcare

Skills

Languages Python, JavaScript, C
Frameworks PyTorch, NumPy, Pandas, Matplotlib, Scikit-learn, React.js

Education

University Of Texas at Arlington Fall 2024–present
M.S. (thesis) Computer Science and Engineering

Sharda University (Greater Noida) 2017–2021
B.Tech Computer Science and Engineering

Research Experience

UT Arlington (Arlington, TX) Fall 2024 – Present
python | *numpy* | *pytorch* | *matplotlib* | *pandas*

Graduate Research Assistant

- Statistical analysis on preterm birth (PTB) for a vulnerable cohort with multi-modal data to study novel interactions between multi-causal effects of PTB. Final model achieves 0.92 AUPRC for PTB prediction task.
- Redcap survey design and development for MIDAS study. To be deployed to collect data for 1000+ patients for sickle cell disease (SCD).
- Performed differential gene expression analysis using CLC Genomics Workbench to identify highly expressed genes in vitamin-deficient vs. sufficient groups within the PTB cohort.
- Trained a medical LLM using knowledge distillation, achieving 60 % size reduction while maintaining comparable accuracy.

Open Source Contributions

Oppia Mar 2020 – Jan 2023

Data and Stability team

- Contributed by resolving more than 20+ issues. **Link**
- Authored Apache beam jobs to maintain and manage inconsistent data and perform large scale data transformations.
- Removed a deeply coupled deprecated component (LogicProof Interaction).
- Contributed in python2 to python3 migrations throughout the project.

Projects

Variational Autoencoder (VAE) for Breast Cancer Slide Compression

- Reproduced and implemented the deep learning pipeline from Nasr et al. (2023), targeting a 1:512 compression ratio on cancer histopathology slides using a VAE architecture.
- Validated that learned latent space representations preserved clinically relevant features, supporting downstream tasks such as cancer subtype identification and slide retrieval.
- Benchmarked reconstruction quality for histopathology and general vision datasets (CIFAR-10), analyzing what image characteristics enable extreme compression in medical imaging.

QuadTree-Accelerated Spatial Simulation with Predator-Prey Dynamics

- Built a 2D simulation engine implementing quadtree spatial partitioning, reducing collision detection complexity from $O(n^2)$ to $O(n \log n)$.
- Applied the same spatial indexing principle relevant to whole-slide image (WSI) analysis, where quadtrees are used to efficiently navigate gigapixel tissue images by region.
- Documented the project in a **technical blog post**; interactive demo available **Link**.

Publications & Manuscripts

[1] **Yumkhaibam, B. S.**, et al. "Stable Pipeline for Multi-Modal Analysis on Preterm Birth for Small Datasets." *In preparation*, 2026.

[2] **Yumkhaibam, B. S.** "Multi-Modal Modeling of Preterm Birth Risk in an Underrepresented Black/African-Ancestry Cohort." M.S. Thesis, University of Texas at Arlington, *In preparation*, 2026.