Sayan Ghosal

CONTACT
INFORMATION

3400 N Charles St,
Baltimore, MD 21218

(443) 531-5268
sghosal3@jhu.edu
Linkedin | Website

PROFESSIONAL
SUMMARY

- Data science researcher with 6+ years of experience in complex, high dimensional, and structured data analysis with an focus on computational biology and genomics.
- Contributions include novel models to embed structured data as graphs to regularize whole-brain whole-genome data for pattern detection and biomarkers selection.
- Developed and built a collaborative network with scientists across the globe, and engaged in multiple collaborative projects (Erdös Number 3) which led to 7 publications in peer-reviewed conferences and journals.
- Passionate about merging the gap between machine learning and the scientific domain by fusing *a priori* scientific information about real-world datasets with data-driven approaches.

EDUCATION

Johns Hopkins University, Ph.D., Summer 2023 (expected)
Department of Electrical And Computer Engineering
Advisor: Prof. Archana Venkataraman

Johns Hopkins University, M.S.E., Spring 2021
Department of Applied Mathematics And Statistics
GPA: 4.0/4.0

Jadavpur University, B.E., 2017
Department of Electronics and Telecommunication Engineering
Honors Thesis: A novel non-rigid registration method for zebrafish larval images
Advisor: Prof. Ananda Shankar Chowdhury
CGPA: 9.3/10.0

RESEARCH
INTERESTS

Disease Association, Multiview Learning, Imaging Genetics, Genomic Association Study, Statistical Genetics, Machine Learning, Graph Learning, Deep Learning

RESEARCH SKILLS

- Predictive modeling.
- Designing robust optimization algorithms.
- Building data modality agnostic transferable and generalizable models.
- Graph representation

- Genetic association analysis.
- Interpretable Deep Learning.
- Incorporating hierarchical structures in deeplearning models.
- Developing interpretable biomarker detection techniques.

WORK
EXPERIENCE

Google X
AI Resident   2022
Project:Integrating Time series and Genetics Data Using Machine Learning To Develop Tools For Sustainable Agriculture.

→ Conceptualized, implemented, and successfully combined traditional statistical models (MLM) with state-of-the-art transformer models to develop sustainable agriculture from genomics and time-series data.

Siemens Healthineers
Machine Learning Intern   2021
Project: Cardiac Outcome Prediction from Multi-Modal Data Leveraging
Prior Population Level Characteristics of the Disease. ISMRM

→ Built a graph-based deep neural network to integrate multimodal data
for risk prediction of cardiovascular disease while exploiting the population
level characterization of the disease.

Department of Computing Science, University of Alberta
Mitacs Globalink Intern   2016
Project: Deformable registration of brain MRI using deep learning. Pattern
Recognition Letters

→ Implemented and published one of the early works of using deep learning
for registration of MRI brain images.

RESEARCH
EXPERIENCE

Johns Hopkins University
Graduate Research Assistant   2017 -
Advisor: Prof. Archana Venkataraman, Department of Electrical and Computer Engineering

BEATRICE: Bayesian Fine-mapping from Summary Data using Deep Variational Inference
- Developed a deep variational approach to find credible sets and posterior inclusion probabilities from GWAS results.
- Utilized a deep neural network as an inference machine to estimate the parameters of our proposal distribution.

GUIDE: A Biologically Interpretable Graph Convolutional Network to Link Genetic Risk Propagations and Imaging Biomarkers of Disease
- Developed a novel deep neural network for whole-brain and whole-genome analysis.
- Tracks the flow of genetic risk through the biological pathways using hierarchical graph convolution and attention operations.
- Fused the imaging and genetic embedding for disease classification.
- Implemented a Bayesian feature selection strategy to extract the discriminative biomarkers of each modality.

G-MIND: An End-to-End Multimodal Imaging-Genetics Framework for Biomarker Identification and Disease Classification
- Developed a novel deep neural network to integrate high dimensional multimodal data like brain imaging and genetics data.
- Performed a classification task, while handling missing data.
- Identified and ranked features using an interpretable network.

Bridging Imaging, Genetics, and Diagnosis in a Coupled Low-Dimensional Framework
- Designed a novel generative-discriminative model for disease classification and feature selection.
- Developed an ADMM based robust optimization strategy that enhances the reproducibility of the model.

- Incorporated biological structures by using the interconnectedness information of different features.

University of Alberta, Canada
Research Assistant   2016
Advisor: Prof. Nilanjan Ray, Department of Computing Science Science.

Deep deformable registration: enhancing accuracy by fully convolutional neural net
- Implemented a deep deformable registration algorithm for MRI images.

- Improved the non-convex optimization of diffeomorphic demons by introducing Fully Convolutional Neural Networks (FCNN).

PUBLICATIONS

S. Ghosal, *et al.*, *BEATRICE: Bayesian Fine-mapping from Summary Data using Deep Variational Inference.*(*In Prep*).

S. Wu, A. Venkataraman, **S. Ghosal**. *GIRUS-net: A Multimodal Deep Learning Model Identifying Imaging and Genetic Biomarkers Linked to Alzheimer's Disease Severity.* Under review for EMBC, 2023.

S. Ghosal, *et al. A Biologically Interpretable Graph Convolutional Network to Link Genetic Risk Pathways and Neuroimaging Markers of Disease.* ICLR: International Conference on Learning Representations, 2022 (**Accepted**). biorXiv.

S. Ghosal, *et al. A Generative Discriminative Framework that Integrates Imaging, Genetic, and Diagnosis into Coupled Low Dimensional Space.* NeuroImage: 238:118200, 2021

S. Ghosal, *et al. G-MIND: An End-to-End Multimodal Imaging-Genetics Framework for Biomarker Identification and Disease Classification.* Proc. SPIE, Medical Imaging 2021: Image Processing. arXiv:2101.11656
Selected for Special Oral Presentation (<15% of Papers), and received best student paper award

S. Ghosal, *et al. Bridging Imaging, Genetics, and Diagnosis in a Coupled Low-dimensional Framework.* MICCAI: Medical Image Computing and Computer Assisted Intervention, 2019. **Selected for Early Acceptance (Top 18% of Submissions)**

S. Ghosal, *et al. A generative-predictive framework to capture altered brain activity in fMRI and its association with genetic risk: application to Schizophrenia.* Proc. SPIE 10949, Medical Imaging 2019: Image Processing.

S. Ghosal, Nilanjan Ray. *Deep deformable registration: Enhancing accuracy by fully convolutional neural net.* Pattern Recognition Letters.

S. Ghosal, *et al. A novel non-rigid registration algorithm for zebrafish larval images.* 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2017.

ADVISING
ACTIVITY

Sarah Wu   2021-Present
Dept: Electrical and Computer Engineering, JHU
Project: Imaging-genetics approaches for Alzheimer's disease prediction.
Current work is under review for EMBC, 2023.

HONORS AND
AWARDS

MINDS Fellow	2022
MICCAI Student Participation Award	2020
Dept. of Electrical and Computer Engineering, JHU, PhD fellowship	2017-2018
Mitacs Globalink Research Fellowship Award	2016

RELEVANT SKILLS   Programming Languages: MATLAB, Python, LaTeX
Libraries: PyTorch, scikit-learn, JAX
Genomic Tools: PLink, RICOPILI