

SAYONI CHAKRABORTY

Dallas, TX | +1 945 264 0363 | scx220042@utdallas.edu | [LinkedIn](#) | [Google Scholar](#)

EDUCATION

PhD, Mathematics, The University of Texas at Dallas

GPA 3.72 | *Expected May 2027*

Research Focus: Machine Learning for Medical Imaging & Biomedical Data Science | Topological Data Analysis | Spatial Omics

M.S., Mathematics, Data Science Specialization, UT Dallas

Graduate Certificate in Data Science

Dec 2025

M.S., Mathematics, Jadavpur University, Kolkata, India

2019 - 2021

B.S., Mathematics, Jadavpur University, Kolkata, India

2016 - 2019

TECHNICAL SKILLS

Programming & Computing: Python, R, SQL, Linux/Unix, HPC & GPU computing | **Data Processing & Analytics:** Data cleaning, preprocessing, validation, data harmonization, exploratory data analysis (EDA), feature engineering, multimodal data integration, reproducible analytical workflows, data visualization | **Machine Learning & Deep Learning:** Classification, Regression, Predictive Modeling, AI Model Development, Random Forests, Gradient Boosting, SVMs, CNNs, Vision Transformers, Multimodal Learning, Contrastive Learning, Ensemble Methods | **Statistical Analysis & Model Evaluation:** Statistical inference, hypothesis testing, regression, cross-validation, bootstrap/resampling, ROC-AUC, PR-AUC, class-imbalance analysis, robustness analysis, model validation, model interpretability | **Biomedical, Omics & Clinical Data:** Spatial Omics, Spatial Proteomics, Tumor Microenvironment Analysis, Multimodal Biomedical Data, Retrospective Clinical Cohorts, Registry & Patient-Level Clinical Data, CT, MRI, X-ray, Digital Pathology, DICOM/NIfTI, Medical Image Preprocessing, Segmentation & Classification | **ML Frameworks & Scientific Libraries:** PyTorch, TensorFlow/Keras, scikit-learn, XGBoost, NumPy, Pandas, SciPy, Matplotlib, pydicom, NiBabel, SimpleITK | **Topological Data Analysis:** giotto-tda, Persistent Homology, Cubical, Point-Cloud & Graph-Based Filtrations, Betti Curves, Persistence Images/Silhouettes, Multiparameter Persistence, Topology-Based Feature Engineering

RESEARCH & PROFESSIONAL EXPERIENCE

Graduate Researcher — MD Anderson Cancer Center, Houston, TX

June 2026 – Present

- Developing a scalable, reusable computational biology framework for high-dimensional spatial omics data, integrating topological methods, statistical analysis, and visualization to characterize tumor microenvironment organization, tissue architecture, and cellular organization across cancer datasets.
- Designing and implementing topology-based representations, including persistent homology and biologically informed filtrations, to quantify cellular heterogeneity, cell-cell spatial relationships, and tumor microenvironment structure across disease stages.
- Building reproducible HPC analytical workflows for data preprocessing, feature extraction, statistical comparisons, robustness evaluation, and visualization; applying the framework to lung adenocarcinoma spatial proteomics data.

Doctoral Researcher — Coskunuzer Lab, UT Dallas

Fall 2022 - Present

Multimodal Clinical AI for Liver Transplantation — *collaboration with University of Arizona College of Medicine*

- Large-scale clinical data integration:** Led the development of an end-to-end reproducible analytical pipeline integrating multi-institutional CT imaging, imaging-derived biomarkers, and OPTN/UNOS clinical registry data from 2,709 deceased donors into an analysis-ready cohort for biopsy-verified liver steatosis prediction.
- Predictive modeling and validation:** Developed and benchmarked multimodal machine-learning models combining clinical variables, CT biomarkers, and topological imaging features, achieving ROC-AUC 0.82 and PR-AUC 0.48 for clinically significant steatosis detection in an imbalanced cohort with 11.2% high-steatosis prevalence.
- Statistical evaluation, interpretation, and scientific communication:** Compared clinical, imaging, topological, and deep-learning models using statistical analysis and temporal holdout validation to assess predictive performance and generalization; interpreted clinically meaningful findings and model limitations and led first-author manuscript preparation with clinical collaborators.

Topology-Informed AI for Cancer Imaging and Biomedical Applications

- AI model development and predictive modeling:** Developed topology-informed CNN and Transformer architectures integrating persistent-homology for high dimensional biomedical imaging data, including 3D MRI, CT, X-ray, dermoscopy, and digital pathology.
- Method development and multimodal learning:** Designed novel sliding-band topological representations, topology-aware contrastive learning, and multiparameter persistence-Transformer frameworks, achieving 10–12 percentage-point improvements in ROC-AUC over strong deep-learning baselines on selected benchmarks.
- Model evaluation, robustness, and research leadership:** Conducted cross-dataset benchmarking, ablation studies, statistical testing, and interpretable feature analysis to identify biologically and structurally meaningful predictive patterns; mentored graduate and undergraduate researchers in developing research projects, experimental design and scientific communication.

PUBLICATIONS

- Chakraborty, S.** et al. A national biopsy-anchored benchmark reveals a CT performance ceiling for donor liver macro steatosis assessment. Under review at American Journal of Transplantation.
- Tastan, F. *, **Chakraborty, S. ***, Lee, S., Coskunuzer, B. [MultiPersistence Topological Fusion with Vision Transformers for Skin Cancer Detection](#). *MIDL 2026*, PMLR 315:2071–2096. (*equal contribution)
- Chakraborty, S.** et al. [Topoformer: Topology-Infused Transformers for Medical Imaging](#). *Machine Learning for Health 2025*, PMLR 297:23–40. Also accepted at **NeurIPS 2026 WiML Workshop**.
- Chakraborty, S.** et al. [Topology-Aware Deep Models for Skin Lesion Classification](#). *ISIC at MICCAI 2025*; **Honorable Mention Award**.
- Fatema, S., Nuwagira, B., **Chakraborty, S.**, Gedik, R., Coskunuzer, B. [TopOC: Topological Deep Learning for Ovarian and Breast Cancer Diagnosis](#). *Topology- and Graph-Informed Imaging Informatics*, Springer LNCS 15239, 2024

AWARDS & HONORS

- **Top 15 Abstract Finalist, AIMed26 Abstract Competition**, Graduate Student Category 2026
- **Mei-Lein Fellowship Award**, School of Natural Sciences and Mathematics, UT Dallas 2026
- **Second Place Poster Award**, ASA North Texas Chapter, DFW Data Science & Statistics Day 2026
- **Honorable Mention Award**, ISIC Workshop at MICCAI 2025
- **Travel Awards**: CHIL 2026; ML4H 2025; DL4Sci Summer School, Lawrence Berkeley National Laboratory 2025; SIAM 2025-2026

INVITED TALKS & PRESENTATIONS

- **Minisymposium Speaker**, SIAM Conference on Imaging Science (IS26) 2026
- **Invited Speaker**, AMS Special Session, George Washington University 2026
- **Invited Speaker**, STDC 2026, Applied Topology and Data Session Mar 2026
- **Poster Presenter**, ASA North Texas Chapter, DFW Data Science & Statistics Day Mar 2026
- **Invited Speaker**, AMS Special Session on Advances in Applied Topology and Topological Data Analysis Fall 2025
- **Invited Speaker & Poster Presenter**, SIAM Texas–Louisiana Section Annual Meeting 2025
- **Accepted Paper**, Women in Machine Learning (WiML) Workshop at **NeurIPS 2026**

SELECTIVE PROGRAMS

- **AxiOM: Machine Learning for Mathematics**, SLMATH (Simons Laufer Mathematical Sciences Institute), Berkeley Mar 2027
- **CHIL Doctoral Symposium**, Conference on Health, Inference, and Learning Jun 2026
- **Foundations of Computational Geometry and Topology**, ICERM, Brown University May 2026
- **DL4Sci Summer School**, Lawrence Berkeley National Laboratory 2025

INDUSTRY PROBLEM-SOLVING WORKSHOPS

- **Mathematical Problems for Industry (MPI)** — Raytheon Technologies, Claremont Jun 2025
Applied CosNet-based ML models and surrogate matrix learning to direction-of-arrival estimation in large-scale signal processing.
- **Mathematical Problems for Industry (MPI)** — Vironix Health, University of Vermont Jun 2024
Performed data analysis and predictive modeling of chronic kidney disease progression.
- **Graduate Student Mathematical Modeling Camp (GSMMC)** — University of Delaware Jun 2024
Built a PDE model of colonic crypt cell-population dynamics with asymmetric cell division; analyzed it analytically and numerically.

LEADERSHIP & SERVICE

- **President**, SIAM Student Chapter (joint chapter across two universities) 2023 - 2025
Organized research, professional-development, and industry events and fostered student mentorship and collaboration.
- **Session Organizer**, AMS Southeastern Sectional Meeting Oct 2025
Organized a special session on Computational Topology and Geometry in Data Science with invited speakers from multiple institutions.
- **Minisymposia Organizer**, SIAM Texas–Louisiana Sectional Annual Meeting Oct 2024
Hosted three minisymposia on geometric and topological approaches in data science and machine learning.

TEACHING

- **Graduate Teaching Assistant**, The University of Texas at Dallas 2022 - Present
Lead problem-solving sessions for Calculus I, Integral Calculus, and Differential Equations, grade undergraduate and graduate coursework including numerical analysis and complex analysis.