

Kasthuri Kannan, PhD

Associate Professor, MD Anderson Cancer Center

✉ kskannan@mdanderson.org ☎ (551) 284 9106 📍 2130 W Holcombe Blvd. Houston. TX 77494

🔗 <https://kannan-kasthuri.github.io/> 🇺🇸 US Citizen 🌐 [linkedin.com/in/kasthuri-kannan-0a05b43](https://www.linkedin.com/in/kasthuri-kannan-0a05b43)

PROFILE

A seasoned data scientist with 17+ experience and proven track record of delivering directed insights, evidenced by quality publications. I excel in managing complex, large scale data analysis and modeling projects. I am recognized for a strong leadership, guiding a multi-institutional team of 6 scholars and 7 faculty members. Actively seeking to leverage my expertise where I can drive innovation and deliver impactful results. I am also a Research Collaborator at Mayo Clinic, AZ where I guide MD PhD students.

PROFESSIONAL EXPERIENCE

MD Anderson Cancer Center

02/2020 – present | Houston, TX

Associate Professor

- Leading graph database development for multi-omics data integration & identifying critical biomarkers.
- Establishing spatial modeling in pathology has to several publications and deeper understanding of the spatial architecture of tumor tissues. Spearheading spatial transcriptomics efforts in glioblastoma.

New York University

11/2013 – 08/2019 | New York, NY

Assistant Professor

- Implemented bioinformatics pipelines, streamlining data processing and analysis workflows, cutting down the time from data collection to actionable insights. Made significant cancer genomics discoveries.
- Directed data science courses cultivating a talent pool equipped with cutting-edge skills in data science and quantitative biology enhancing NYU's capabilities in data-driven decision-making and innovation.

Memorial Sloan-Kettering Cancer Center

04/2011 – 10/2013 | New York, NY

Research Fellow/Associate

- Established mutation pipeline for Brain, Head & Neck cancers resulting in improved understanding of cancer genetics and tumor heterogeneity. Provided directed insights in various cancer studies.
- Discovered ATRX mutations in lower grade gliomas establishing clinical diagnosis and contributed to a better understanding of glioma biology.

Pennsylvania State University

10/2010 – 03/2011 | State College, PA

Research Associate

- Offered bioinformatics consultation to researchers resulting in the appropriate use of computational tools and statistics to inform actionable insights.
- Managed sequencing tasks, ensuring the smooth operation of sequencing projects and delivering high-quality data for downstream analysis and research.

Stowers Institute for Medical Research

01/2008 – 09/2010 | Kansas City, MO

Research Specialist

- Developed image processing methods for worm/fly tracking leading to publications (*Cell*, *PLoS Genetics*).
- Implemented an automated workflow to process cell images, reducing image acquisition times five-fold.

Knowledge Based Systems, Inc

01/2007 – 12/2008 | College Station, TX

Internship

- Delivered data-driven insights for aircraft movement operations - improved the efficiency and effectiveness of land based air traffic control management at the Tinker Air Force Base, OK.
- Proposed cost and time-saving measures for managing Air Force logistics and identified opportunities to optimize logistics operations, resulting in significant cost savings and improved resource utilization.

ACTIVE DATA SCIENCE PROJECTS

Geospatial modeling of tissues — Spatial point processes are robust statistical tools for analyzing point patterns. By representing cells as points and annotating their measurements (e.g., single-cell gene expression), we can effectively model cell-cell interactions. We leverage the R package *spatstat* to model these interactions and gain insights into cancer from immuno-oncology standpoint.

Biomarker discovery using graph database — Graph databases efficiently represent complex biological networks, facilitating biomarker identification. They enable powerful queries and community detection algorithms, making it easier to explore relationships between multiple genes. Neo4j, a property graph database, along with its Graph Data Science Library, is used to derive insights and propose actionable biological targets.

Biomarker validation using Graph Neural Networks (GNNs) — Insilico biomarker validation is vital for identifying actionable targets for experiments and prognosis. Using GNNs we validate biomarkers from graph databases by leveraging biological network structures like gene-gene interactions. GNNs aggregate neighbor information to learn meaningful representations, capturing local and global patterns for improved gene-disease relationship predictions.

DATA SCIENCE EDUCATION

Developed and taught Programming for Data Analysis, Machine Learning & AI, and Methods in Quantitative Biology courses. Please refer <https://kannan-kasthuri.github.io/#about> for details.

SKILLS

Leadership

Managing data science and bioinformatics projects ● ● ● ● ●

Software (Representative)

Python, R, Java, Cypher, SQL, Bioinformatics tools, Unix/HPC/Bash, Neo4j, MySQL, PostgreSQL, HTML, Javascript, Pytorch, Tensorflow, LangChain ● ● ● ● ●

Data Science/Machine Learning/AI

Math, CS, Statistics, Programming, GenAI, LLMs, Embeddings, RAG/GraphRAG, Ethical hacking ● ● ● ● ●

Bioinformatics & Image Processing

Pipeline development and data engineering
Geospatial analysis ● ● ● ● ●

EDUCATION

Doctor of Philosophy (Computer Science)

Texas A&M University

2002 – 2008 | College Station, Texas

Master of Science (Mathematics)

Texas A&M University

2000 – 2002 | College Station, Texas

Master of Science (Mathematics)

Indian Institute of Technology, Madras

1998 – 2000 | Chennai, India

Bachelor of Science (Mathematics)

University of Madras

1995 – 1998 | Chennai, India

INVITED TALKS (REPRESENTATIVE)

Mayo Clinic (2023), National University of Singapore (2022), Texas A&M University (2021), Courant Institute of Mathematical Sciences (2019), Institute of Mathematical Sciences, w/ honorarium (2018)

PUBLICATIONS

Authored/co-authored 33 peer reviewed articles, that includes very high-impact journals. Please refer <https://kannan-kasthuri.github.io/publications/publications.html>