

# SUDARSHAN ANAND

Atlanta, GA (willing to relocate) | [asudarshan14@gmail.com](mailto:asudarshan14@gmail.com) | +1 (404) 453-5664 | LinkedIn: [www.linkedin.com/in/sanand315/](https://www.linkedin.com/in/sanand315/)

Portfolio: [sudarshananand14.vercel.app](https://sudarshananand14.vercel.app) | F-1 OPT (2026–2029)

## SUMMARY

**NSF-EMBS-Google NextGen Scholar and AI Researcher** specializing in the architecture and deployment of multimodal and time-series ML systems. Proven expertise in modeling neurodegenerative disease progression and lead-architecting large-scale analysis of 1.2M+ unstructured public health records. Develops evaluation-driven, robustness-focused frameworks designed to solve extreme metadata sparsity and noise in real-world clinical data. Bridging the gap between novel algorithmic research and scalable, policy-relevant health engineering.

**Interested Roles:** AI/ML Research Scientist, Machine Learning Engineer, or Data Scientist (R&D).

## EDUCATION

### Georgia Institute of Technology

Atlanta, GA, USA

*Master of Science in Computational Science and Engineering (CGPA: 4.0/4.0)*

May 2026

**Relevant Coursework:** Data Science for Epidemiology, Machine Learning for Graphs, Industrial Math, Numerical Linear Algebra

### Birla Institute of Technology and Science (BITS Pilani)

Pilani, Rajasthan, India

*Master of Science in Mathematics and Bachelor of Engineering in Computer Science (CGPA: 9.4/10)*

Aug 2024

## HONORS & AWARDS

### IEEE BHI 2025 Data Challenge Competition Winner

Oct 2025

- Led development of an **AI/ML-based depression risk prediction system**, contributing to a winning solution in the IEEE-sponsored competition (supported by NSF and Google)

### IEEE BHI 2025 Young Professional NextGen Scholar (NSF, EMBS, Google Sponsored)

Sep 2025

- Selected for a **competitive NSF/Google-sponsored scholarship** recognizing high-impact biomedical AI research; invited to present at IEEE BHI 2025

## SELECTED PUBLICATIONS

### Foundation Models / ML Systems

- H. Kamarthi, S. Li\*, **S. Anand\***, S. Das, A. Ramesvaran, A. Bhalerao and B.A. Prakash, “*Samay: A Unified Pipeline Framework for Time-Series Foundation Models*” VLDB 2026 (Demo track, submitted)

### Public Health / Data-Centric AI

- Vasan, S.S.; **Anand, S.**; Lee, M.; Fluck, N.C. On Burden of Diseases, Prevention, Medical Research and Health Service Delivery: Grampian Case Study. *Int. J. Environ. Res. Public Health* 2026, 23, 763.
- **S. Anand**, M. Lee, and S. S. Vasan, “*Metadata makes Big Data useful for AI/ML-ready, equitable pandemic response*” IEEE-EMBS BHI 2025 (Poster)
- **S. Anand**, M. Lee, and S. S. Vasan, “*Why de-identified patient meta-data matters during pandemics*” Scotland's Health Research and Innovation Conference 2025 (Poster)

### Optimization / Systems (Deployed Academic Product)

- Shiksha, **Anand S**, Shekhawat K, Agrawal K. *Automated generation of circulations within a floorplan. Artificial Intelligence for Engineering Design, Analysis and Manufacturing*. 2025.

## RESEARCH & PROGRAMMING PROJECTS

### DRiP: White-box framework for Depression Risk Prediction

Oct 2025

*IEEE-EMBS International Conference on Biomedical Health Informatics (BHI) 2025*

Atlanta, GA, USA

- Built **interpretable longitudinal depression risk model for sparse clinical data (n=200)**, enabling feature-level temporal attribution under constrained model complexity
- Developed **stacked classical ML ensemble** achieving  $R^2 = 0.80$  (vs.  $\sim 0.40$  baseline) on patient-level held-out real data under 10-fold cross-validation
- Evaluated distribution-matched synthetic augmentation ( $\sim 600$  samples) and performed post-hoc temporal attribution to extract stable risk and treatment-response signals.

### Multimodal Progression Tracking of Neurodegenerative disorders

Aug 2025 – May 2026

*Georgia Institute of Technology*

Atlanta, GA, USA

- Built a multimodal pipeline for Parkinson's disease progression modeling, fine-tuning a **SimCLR-based self-supervised model (ResNet-18 backbone)** on 3-class staging (control vs prodromal vs Parkinson's) using structural MRI with demographic fusion (age, gender) for holistic patient information.
- Evaluated a clinically ambiguous classification task with **weak structural separability between early disease stages**, using class-wise F1 (macro F1  $\approx 0.60$ ) to better reflect overlapping phenotypes instead of accuracy.
- Scaled **distributed PyTorch training on HPC clusters using SLURM** for large longitudinal neuroimaging workloads.
- Developed and debugged HPC infrastructure workflows using **LLM-assisted engineering (Claude-based tooling)** to accelerate deployment and scaling of MRI processing pipelines.

## Investigating importance of Patient Metadata in Public Health

Edith Cowan University

May 2025 – present  
Western Australia, Australia

- Led analysis of **1.2M unstructured COVID-19 patient records** to evaluate patient metadata utility for modeling host–pathogen interactions in resource-constrained public health settings
- Quantified extreme data limitations, identifying a **94% metadata sparsity rate**, identifying missing metadata as a primary bottleneck for reliable resource planning and clinical interventions during pandemic response.
- Used **state-of-the-art LLMs (Gemma-2, MedGemma-1.5)** as a controlled probe to evaluate robustness under degraded clinical context and missing information regimes.

## Samay: Time-series Foundational Models Library

Georgia Institute of Technology

Jan 2025 – May 2025

Atlanta, GA, USA

- Developed **Samay**, an open-source library **unifying 10 time-series foundation models** for forecasting, classification, and anomaly detection, enabling standardized inference and evaluation workflows (**34 stars, 7 forks**)
- Developed modular **PyTorch-based inference and fine-tuning pipelines (PyTorch 2.0+)**, enabling zero-shot evaluation and task-specific adaptation of pretrained models across diverse time-series formats.
- Designed and executed large-scale benchmarking against the **GiftEval benchmark**, validating model performance and scalability across unified evaluation protocols for time-series foundation models.

## Viral evolution in heterogeneous human contact networks

Georgia Institute of Technology

Oct 2024 – Dec 2024

Atlanta, GA, USA

- Developed a multi-strain epidemiological framework (SI<sub>1</sub>I<sub>2</sub>I<sub>3</sub>SD) to simulate viral competition and mutation-driven strain evolution across populations.
- Analysed how network heterogeneity (community structure and long-range links) shapes inter-strain competition and outbreak dynamics in evolving contact networks. [[project website](#)][[Github](#)]

## TEACHING EXPERIENCE

### Graduate Teaching Assistant

Georgia Institute of Technology (CSE Algorithms (Aug-Dec '25); Numerical Linear Algebra (Jan '26-May '26))

Aug 2025 – May 2026

Atlanta, GA, USA

- Facilitate instruction for **180 students in both Algorithms and Numerical Linear Algebra**, providing targeted conceptual reinforcement and assignment feedback

## PROFESSIONAL EXPERIENCE

### AI Product Development Intern

Rezolve.ai

Jun 2025 – Aug 2025

Dublin, CA, USA

- Built an enterprise **Retrieval-Augmented Generation (RAG) search MVP**, for interacting with company policy documents, using FastAPI, Pinecone, and PostgreSQL with conversational memory, enabling policy retrieval in **<20s latency** for enterprise knowledge workflows.
- Developed a **prefix-tree (trie)-based autocomplete system** and automated global data pipeline updates using **GCP Cloud Run**.
- Designed prototype for an **Agentic AI ChatOps system (Slack/Teams)** for IT incident response, automating alert triaging and runbook generation to streamline incident workflows.

### AI Scientist Intern

Qure.ai Technologies Pvt. Ltd.

Jan 2024 – Jun 2024

Bangalore, Karnataka, India

- Improved a clinical lung nodule detection system, increasing **Pearson correlation with radiologist assessments by ~45% relative to baseline models**, enhancing reliability for early cancer screening workflows.
- Benchmarked image registration performance for longitudinal CT-based lung nodule tracking, evaluating **temporal alignment consistency across consecutive scans** for clinical follow-up analysis.

## TECHNICAL SKILLS

**Programming & Data:** Python, SQL, Pandas, NumPy, PostgreSQL, Neo4j, Git, Linux/Unix

**Machine Learning & AI:** PyTorch, Unsloth, multimodal learning, computer vision, self-supervised learning, LLM applications, retrieval-augmented generation (RAG), prompt engineering, time-series foundation models, transformer architectures, model fine-tuning, model evaluation, inference optimization, feature engineering, A/B testing, statistical modeling

**Systems & Infrastructure:** FastAPI, Pinecone, ChromaDB, GCP, SLURM, scalable ML pipelines, cloud deployment, API development, backend systems, ML workflow automation, CI/CD, production ML systems

**Data Science & Analytics:** Data preprocessing, exploratory data analysis (EDA), performance optimization, predictive modeling, longitudinal data analysis, unstructured data processing

**Professional:** Cross-functional collaboration, stakeholder-facing collaboration, technical documentation, research communication, project leadership