

Dhananjay Balakrishnan

University Email: dhananjb@stanford.edu | [Personal Webpage](#)

EDUCATION

Stanford University

Sept 2024 - June 2026

MS in Computational & Mathematical Engineering - GPA: 3.86/4.00

Stanford, CA, USA

- Relevant Coursework: *Theory of Stats, Introduction to AI Governance, Intro to Causal Inference, Geometric and Topological Data Analysis, Introduction to Parallel Computing, Statistical Consulting Workshop, Introduction to GIS, Mining Massive Datasets*

Indian Institute of Technology, Madras

July 2019 - May 2024

Inter-disciplinary Dual Degree in Mechanical Engg. and Data Science - CGPA: 9.17/10.0

Chennai, India

- Minor in **Mathematics**
- Thesis: A Self-Supervised Approach to Motion Deblurring. Guided by [Dr. AN Rajagopalan](#). [\[Report Link\]](#)
- Awarded the IIT Madras **Governor's Gold Medal**, for **best outgoing dual-degree student**.
- Received the IIT Madras **Institute Blues Certificate of Merit**, given to outstanding graduating students.
- Relevant Coursework: *Pattern Recognition & Machine Learning, Reinforcement Learning, Multi-Armed Bandits, Modern Computer Vision, Big Data Laboratory, Fundamentals of Deep Learning, Numerical Optimization, Game Theory*

EXPERIENCE

Data Science Intern, STAA (Science, Technology Assessment, and Analytics)

January 2026 - May 2026

Innovation Lab, GAO (Government Accountability Office)

Washington DC, USA

- Worked in a **cross-functional team** to design and implement a **scalable LLM-based concept induction pipeline** to assist analysts in processing large volumes of **internal audit documents**, directly impacting **300+ stakeholders** across the GAO.

Junior Machine Learning Engineer, YUX Cultural AI Lab

June 2025 - August 2025

Stanford SEED Internship [\[Work Portfolio\]](#) Presented at AfricaNLP Workshop, [EACL 2026](#), Rabat.

Dakar, Senegal

- Advised on **AI localization efforts** for **indigenous, Afro-centered contexts** at YUX, a Senegal-based research firm.
- Worked on **context-specific speech transcription** models for **low-resource languages** such as Wolof and Hausa.
- Built an **African stereotypes dataset** from 1,000+ survey responses; evaluated **Afro-specific bias** in open-source LLMs.

Student Researcher, World Bank Health Group

March 2025 - June 2025

CME291: ICME Xplore, World Bank [\[Report Link\]](#) Presented at [SimBig 2025](#), Lima, Peru.

Stanford, CA, USA

- Collaborated in a 3-member team with the **World Bank Health Group** to develop a **geospatial framework** integrating **clinic locations, road networks, and internet connectivity** to assess **healthcare accessibility** across South Africa.
- Identified underserved communities and **healthcare accessibility gaps**, supporting **equitable infrastructure investment** and resource allocation through geospatial analysis

Commodities Strats and Quant Intern, Fixed Income Division

May 2023 - August 2023

Morgan Stanley

Mumbai, India

- Implemented various mathematical models associated with volatility & risk, option pricing, and Greek calculation in C++.
- Helped migrate from the **JWSS7** to the **CiQuick** model to better fit the **volatility smiles** data, leading to an improvement of $\sim 7\%$ in the explained PnL (Profit & Loss). This has been **pushed into production**.

Project Associate, Crack Detection and Segmentation

June 2021 - October 2022

Computer Vision Club, IIT Madras. Presented at the CVCIE Workshop, ECCV 2022.

Chennai, India

- Compiled a uniformly annotated **open-source dataset** with **9000+ image-mask pairs** for the crack segmentation task.
- Used **ViT-generated features** to improve the generalizability of the model, further improving the MIOU by $\sim 5\%$.

ACADEMIC PROJECTS & COMPETITIONS

Analyzing Caste-Based Child Nutrition Disparities Using NFHS-5 Data

September 2025 - November 2025

Guide: [Dominik Rothenhäusler](#). Course: STATS209 - Introduction to Causal Inference. [\[Report Link\]](#)

- Applied **difference-in-differences** and **triple-difference** models to evaluate the impact of COVID-19 on child nutrition outcomes across caste groups using India's NFHS-5 survey data.
- Identified **persistent disparities in child malnutrition**, with marginalized caste groups experiencing $\sim 4.2\%$ **higher rates of stunting and underweight**, highlighting **structural inequities in health outcomes**.

Mapping Socio-Ecological Changes in the Red Corridor using GIS

October 2025 - November 2025

Guide: [Diana Moanga](#). Course: EARTHSYS 144 - Introduction to GIS [\[Report Link\]](#)

- Used **night-time light** and **soil-cover** remote sensing data to measure **indigenous settlement displacement** and **diverging ecological and developmental outcomes** around iron and bauxite mining sites in India's Red Corridor.

Evaluating Structural Discrimination in Indian AI Systems

March 2025

Guide: [Anka Reuel](#). Course: CS134 - Introduction to AI Governance. [\[Report Link\]](#)

- Used **textual embeddings** to evaluate the **caste bias** of state-of-the-art **multilingual LLMs**, measured across various Indian regional languages. Further, performed a resume audit experiment to study **structural biases** in LLM-based decision making.

Stochastic Modelling of Gene Regulatory Networks

May 2024

Guide: [Dr Karthik Raman](#). Course: BT5240 - Computational Systems Biology. [\[Report Link\]](#)

- Used the **Gillespie Algorithm** and **Stochastic Boolean Networks (SBN)** to model the **Lac Operon** gene network.
- Introduced a novel **transitory-state** based approach to incorporate **history dependence** into the above SBN model.
- Proposed a **Probability Search Algorithm** coupled with a **custom loss function**, to accurately model the system.

Modelling Neuronal Dynamics as a Random Process

May 2023

Guide: [Dr Sayan Gupta](#). Course: AM5340 - Stochastic Processes in Mechanics [\[Report Link\]](#)

- Numerically simulated various time histories of the state variables using **Shinozuka's Spectral Decomposition** method.
- Developed a **Polynomial Chaos Expansion** model for the state variable, and used it to **validate** the MC Simulations.
- Carried out a **parametric study** on the effect of varying the noise σ on the **exceedance probability** of the response.

Indic Language Transliteration using Seq2Seq Models

May 2023

Guide: [Dr Mitesh Khapra](#). Course: CS6910 - Fundamentals of Deep Learning.

- Built an **English-Tamil transliteration** model, using an RNN model with attention to achieve a validation accuracy of **0.4751**.
- Used the evolution of model predictions to make **qualitative observations** on the **semantics** of the Tamil script.

Bronze Medal, Super Resolution of Lunar Surface Images

January 2023

Inter-IIT Tech Meet 2023

Kanpur, India

- Achieved a PSNR of 22 by using an SRGAN model to perform a 16x super-resolution task on lunar surface images, provided by ISRO. Based on the performance of our model, we were awarded the **bronze medal**, from over **25 teams**.

EXTRA-CURRICULAR ACTIVITIES

- **Teaching Assistant:** Partial Differential Equations in Engineering, Stanford - Class of 43 students taught by Prof. Sanjiva K Lele.
- **Quizzing:** Active member of the **IIT Madras Quiz Contingent** from 2019-24, part of the **winning contingent** at [Nihilanth](#) 2020 & 2023. Part of the quiz contingent that won the overall quiz cup at the Inter-IIT Cultural Meet 2023 & 2024.
- **Student Mentorship:** Mentored underprivileged students from JNV Pondicherry to prepare for the JEE Main and Advanced examinations, as part of the Avanti Fellows Pondicherry Chapter.

PUBLICATIONS

- Le Beux, Y., Audu, O., Ankeli, O. D., **Balakrishnan, D.**, Weya, M., Ralaifarinosy, M. D., & Ezeani, I. (2026, May). AfriStereo: A Culturally Grounded Dataset for Evaluating Stereotypical Bias in Large Language Models. In Proceedings of the Fifteenth Language Resources and Evaluation Conference (LREC 2026) (Vol. 11, No. 16, pp. 266-280).
- Basilwango, E., Le Beux, Y., Ankeli, O. D., Berdys, P. H., & **Balakrishnan, D.** (2026, March). Enhancing Automatic Speech Recognition Models for Maternal and Reproductive Health: Fine-Tuning and Real-World Evaluation in Wolof. In Proceedings of the 7th Workshop on African Natural Language Processing (AfricaNLP 2026) (pp. 256-263).
- **Balakrishnan, D.**, Drekshagen, S., Sahu, V., Tariverdi, M., & Nunez-del-Prado, M. (2025, October). Measuring Healthcare Accessibility Through Clustering Techniques and Multi-layer Network Analysis. In Annual International Conference on Information Management and Big Data (pp. 109-124). Cham: Springer Nature Switzerland.
- Kulkarni, S., Singh, S., **Balakrishnan, D.**, Sharma, S., Devunuri, S., & Korlapati, S. C. R. (2022, October). CrackSeg9k: A collection and benchmark for crack segmentation datasets and frameworks. In European conference on computer vision (pp. 179-195). Cham: Springer Nature Switzerland.

DEVELOPER TOOLS AND LANGUAGES

Frameworks & Software: Python, SQL, C, C++, MATLAB, Wolfram, Julia, JavaScript, PyTorch, Tensorflow, Keras, NumPy, Sci-kit Learn, Open-CV, SciPy, Matplotlib, MS Excel, GNU Octave, LATEX, ABAQUS, AutoCAD, Simulink, Fusion 360, ArcGIS

Languages: English: Native, Tamil: Native, Malayalam: Native, Hindi: Working Proficiency. **Work Permit:** US Citizen.