Progyan Das

progyan.das@iitgn.ac.in | progyan.me

+91 60098 84968

github.com/plugyawn

Education

Indian Institute of Technology (IIT) Gandhinagar | Bachelors of Technology [2020-2024]
Major: Computer Science and Engineering (Honors), Minor: Physics
GPA: **9.62**/10, Institute Rank **1 of 245**

Scholastic Achievements & Awards

- Presented with the Award for Academic Excellence for highest GPA in the batch for three consecutive AYs: 2021-22, 2022-23, and 2023-24 at IIT Gandhinagar.
- Among 10 students across India awarded Cargill Global Scholarship for outstanding academics and leadership, with 2 years of mentorship and scholarship of 5000 USD. [2022]
- Selected for and attended Cornell, Max-Planck, Maryland Pre-Doctoral Summer School at Max Planck Institute for Software Systems, Saarbrücken, Germany. [2023]
- Caltech Summer Research Fellow (SURF), under Dr. Ashish Mahabal, Centre for Data-Driven Discovery, Caltech and Dr. Nitin Singh, Jet Propulsion Laboratory, NASA. [2023]
- Selected for MITACS Globalink research fellowship, under Prof. Oleksandr Romanko, University of Toronto. [2023]
- Presented with the prestigious Chetan Dhande Scholarship for Academic and Overall Excellence at IIT Gandhinagar. [2022]
- Participated in Google Summer of Code 2022 with the Zulip Organization (resigned after mid-term evaluation to focus on research). [2022]
- Awarded Academic Research Ranking 1 in the Computer Science and Engineering discipline for recommendation for institute-led academic internship and exchange programmes.
- Elected Technical Secretary, Student Council by electoral pool of 2000+ students of the IIT Gandhinagar community. [2023-24]

Publications & Research

1. Progyan Das*, Dwip D. Dalal*, Anirban Dasgupta, *ODE-Solvers are also Wayfinders: Neural ODEs for Multi-Agent Path-Planning*, **accepted** in Symbiosis of Deep Learning and Differential Equations @ NeurIPS 2023, under preparation for ICML 2024
2. Mihir Agarwal*, Progyan Das*, Udit Bhatia, *Spatially-Regularized Graph Attention Autoencoders for Rainfall Anomaly Detection*, **accepted** in Climate AI @ NeurIPS 2023
3. Progyan Das*, Mihir Agarwal*, *Less But Better: Towards better AQ Monitoring by learning Inducing Points for Multi-Task Gaussian Processes*, **accepted** in Adaptive Experimental Design and Active Learning @ NeurIPS 2023
4. Progyan Das*, Mihir Agarwal*, *Focus on What's Important! Inspecting Variational Distributions for Gaussian Processes for better AQ Station Deployment*, **accepted** in Computational Sustainability @ NeurIPS 2023, under preparation for UAI 2024
5. Aalok Gangopadhyay*, Dwip D. Dalal*, Progyan Das*, Shanmuganathan Raman, *Flow Symmetrization for Parameterized Constrained Diffeomorphisms*, in review at Eurographics
6. Dwip D. Dalal*, Progyan Das*, Mihir Agarwal*, Soumyabrata Chakrabarty, Dibyendu Chakrabarty, *Uncertainty Bounds for Anomalous Geomagnetic Storm Forecasting*, in review at Advances in Space Research

Under Preparation

7. Progyan Das*, Navneet Bung, *Disentanging the Peptide Space: Contrastive Variational Autoencoders for Generating Antimicrobial Peptides*, submitted to PLOS Biology.
8. Progyan Das*, Ashish Mahabal, *Fourier Convolutional Transformers for irregular light-curve time-series classification*, currently under preparation.
9. Progyan Das*, Ashish Mahabal*, Nitin Singh*, *PredictContamination: World's Biggest Bacterial Property dataset*, currently under preparation.

A rolling list with links to all projects and publications available on my website.

Relevant Courses Undertaken

Core Computer Science Courses

Operating Systems `[A]`, **Introduction to Data Science** `[A]`, **DSA II (Algorithms Design)** `[A]`, Compilers `[A-]`, Discrete Mathematics, Computer Architecture and Organization, Networks and Information Security†

Graduate Computer Science Courses

Optimization Methods for Machine Learning `[A]`, Natural Language Processing `[A]`, Probabilistic Machine Learning, Understanding LLMs, Advanced Deep Learning†, Advanced Algorithms†

Mathematics, Statistics and Physics

Linear Algebra & Calculus (Mathematics I) `[A+]` (`11/10`), Multivariate Calculus & Complex Analysis (Mathematics II) `[A]`, Differential Equations (Mathematics III) `[A]`, **Networks and Complex Systems** `[A]`, **Inverse Modelling in Physical Sciences** `[A]`, Electromagnetism & Quantum Mechanics (Physics I) `[A]`, Computational Physics†

† indicates courses for upcoming semester. `A` is 10/10, `A+` is 11/10, awarded in exceptional cases.

Exceptionally relevant courses in **bold**.

Research Experience

California Institute of Technology (Caltech) | Research Fellow [2022-23]

Irregular Time-Series Modelling for Light Curve Representation and Classification

Supervisor: *Prof. Ashish Mahabal (Caltech)* Summer 2023 onwards

- Devised the first deep learning pipeline for raw light-curve classification by combining convolutional transformers with contrastive learning, achieving competitive performance with SOTA with less memory footprint in comparison to Neural ODE based methods.
- By combining state-of-the-art vision models with wavelet transform, beat SOTA accuracy on light-curve classification, up to 82.8% over 10 classes from 76.8% over 7 classes previously.

Machine Learning over tabular datasets for Bacterial Contamination on spacecrafts

Supervisor: *Prof. Ashish Mahabal (Caltech), Dr. Nitin Singh (JPL)* Winter 2022, Spring 2023

- Curated, pre-processed and modularized the world's largest (to the best of our knowledge) dataset of bacterial information, wrote a terminal user interface for researchers to dynamically update the dataset (of 160,000+ bacteria) from BacDive and run analyses over the data.
- Established new SOTA benchmarks in predicting bacterial properties through tree-based learning methods, contrasted with relevant deep learning architectures (eg. Tabformer).

Data Science Lab, IIT Gandhinagar, under *Prof. Anirban Dasgupta* [2023]

On the scalability of Temporal Graph Neural Networks

Supervisor: *Shubhajit Roy, Prof. Anirban Dasgupta* Spring 2023, Fall 2023

- Researched sketch-based methods, dependencies and mathematical guarantees for sparser graphs, for enchanced scalability for Temporal Graph Neural Networks, over a novel network construction over the ERA5 dataset based on inputs from climate and network scientists. **Part of this work was accepted to the** ClimateAI **workshop at** NeurIPS'23.

Neural Ordinary Differential Equations for Solving Mazes

Supervisor: *Dr. Aalok Gangopadhyay (TIFR), Prof. Anirban Dasgupta* Spring 2023, Fall 2023

- Ideated, mathematically derived, and coded a novel path-finding algorithm based on masked Neural ODEs. Our method relies on the non-intersection guarantees of Neural ODE field lines along with a novel flow-masking mechanism to solve multi-agent pathfinding problems. **Part of this work was accepted to the** Symbiosis of Deep Learning and Differential Equations **workshop at** NeurIPS'23. Full work in preparation for ICML'24.

Private Coresets for Neural Networks through NTK-based Gaussian Processes

Supervisor: *Shrutimoy Das, Prof. Anirban Dasgupta* Spring 2023, Fall 2023

- Came up with a method for constructing architecture-dependent private coresets by Gaussian Process with a Neural Tangent Kernel, that performs better than all standard baselines. **Under preparation for** Transactions in Machine Learning Research.

Uncertainty-quantifying Neural Operators with Gaussian Process kernel

SUPERVISOR: *Prof. Anirban Dasgupta* Spring 2023, Fall 2023

- Integrated a sparse variational Gaussian Process equipped with differentiable kernel composability into the kernel integral in the Neural Operator framework, to learn uncertainty bounds for operator predictions with universal approximation guarantees.

CVIG LAB, IIT GANDHINAGAR under *Prof. Shanmuganathan Raman* [2022-23]

Flow Symmetrization for Parameterized Constrained Diffeomorphisms

SUPERVISOR: *Dr. Aalok Gangopadhyay, Tata Institute of Fundamental Research (TIFR)*

- Co-developed FLOW SYMMETRIZATION, a novel method for learning symmetrically constrained flows via incorporating fourier features in the Neural ODE weights, with massive applications over multiple fields, such as efficiently learning distributions over topological manifolds like spheres and tori, or for differentiable shape optimization.
- A part of FLOW SYMMETRIZATION can be used for solving the famous Escherization problem — find an shape that tiles the R^2 plane which closely approximates a given proposal shape. By leveraging identification topology, it can be used to learn constrained diffeomorphisms for the plane that result in unconstrained diffeomorphisms on the spheres, tori, et cetera. **Part of this work is currently under review in** EUROGRAPHICS 2023.

MACHINE INTELLIGENCE AND RESILIENCE LAB under *Prof. Udit Bhatia* [2022-23]

Geometric Deep Learning over Event Synchronization Climate Networks

SUPERVISOR: *Saarth Dubey, Danish Mansoor, Prof. Udit Bhatia* Spring 2023, Fall 2023

- Formulated a Graph Attention Network with a novel loss function for predicting anomalous rainfall regions on an event-synchronization based network. Our work on anomaly prediction for precipitation achieved competitive results at huge speedups compared to classical techniques. **Accepted to the** CLIMATEAI **workshop at** NEURIPS 2023.

Methods for measuring Multivariate Synchrony in Climate Networks

SUPERVISOR: *Saarth Dubey, Danish Mansoor, Prof. Udit Bhatia* Spring 2023, Fall 2023

- Formulated an empirically-motivated, theoretically sound metric involving the event synchrony and the covariance across multiple variables for building better climate networks for data with multiple covariates, with a focus on the publicly available ERA5 dataset.

SUSTAINABILITY LAB, IIT GANDHINAGAR under *Prof. Nipun Batra* [2022-23]

Scalability of Gaussian Process via stochastic variational inference for Air Quality

SUPERVISOR: *Zeel Patel, Prof. Nipun Batra* Spring 2023, Fall 2023

- Thoroughly researched, implemented, and built benchmarks for various Bayesian learning techniques for uncertainty-bound AQ inference, with a focus on scalable multi-task Gaussian processes as part of the SUSTAINABILITY LAB. Our work resulted in four publications, of which I authored two — **accepted at the** COMPUTATIONAL SUSTAINABILITY **and** ADAPTIVE EXPERIMENTAL DESIGN AND ACTIVE LEARNING **workshops at** NEURIPS'23.
- Designed a differentiable kernel composition mechanism for joint learning of kernels in multi-task Gaussian processes via a Neural Kernel Network over output covariance function, significantly improving model fit with negligible effect on training time.
- Built two libraries, `WimpyML` and `GPZoo`, containing from-scratch implementations of various machine learning and Gaussian process algorithms, with a convenient API.

SUMMER RESEARCH FELLOW, IIT GANDHINAGAR under *Prof. Neeldhara Misra* [2022]

Algorithms for Efficient Nearly-Fair Division with Constraints

SUPERVISOR: *Aditi Sethia, Prof. Neeldhara Misra* Spring 2023, Fall 2023

- Researched Nearly-Fair Division algorithms, in an attempt to find polynomial-time solutions for constrained scenarios that is envy-free upto one item, across various efficiency criteria.

TCS RESEARCH | Research Intern [2023]

Generative Modelling of Antimicrobial Peptides via Contrastive Learning

SUPERVISOR: *Dr. Navneet Bung, Prof. Arijit Roy* Spring 2023, Fall 2023

- Implemented the SOTA Wasserstein Autoencoder-based architecture and improved upon it specifically for antimicrobial peptides with contrastive learning for sampling high-quality candidate peptides. **Manuscript to be submitted to** PLOS BIOLOGY journal.

Academic Service & Teaching

- Lead Instructor, SC336: Mathematics of Machine Learning, first completely undergraduate-run credited course offered at IIT Gandhinagar, advised by Prof. Anirban Dasgupta. 170+ students, including undergraduates, postgraduates, and postdocs from various fields.
- Teaching Assistant, Operating Systems | Prof. Rajat Moona (Director, IIT Gandhinagar), Prof. Abhishek Bichhawat, *Fall 2023*
- Teaching Assistant, Natural Language Processing | Prof. Mayank Singh, *Fall 2023*
- Teaching Assistant, Data Structures & Algorithms I | Prof. Neeldhara Misra, *Fall 2022*
- Teaching Assistant, Data Structures & Algorithms II | Prof. Bireswar Das, *Spring 2024*
- Teaching Assistant, Introduction to Data Science |Prof. Anirban Dasgupta, *Spring 2024*
- Reviewer, Symbiosis of Deep Learning and Differential Equations @ NeurIPS'23

As TA, I have graded exam papers, set assignments, and taught tutorial classes for 150+ students.

Industrial Experience

Mudrex Capital | Quantitative Research Intern [Fall 2022]

Stochastic Differential Equations for modelling Highly Volatile Markets

- Researched and implemented Jump Diffusion and other stochastic PDEs for capturing stochasticity for statistical arbitrage, with a focus on highly volatile cryptocurrency markets.
- Used *Latent Stochastic Differential Equations* through `torchsde` to build an autoregressive deep learning pipeline for accurate, uncertainty-quantified time-series forecasting.
- Built an open-sourced python library, `coinbase`, for backtesting of alphas on cryptocurrency markets through an API for implementing and testing various machine learning algorithms.

Zulip Organization | Google Summer of Code Fellow [2023]

UI/UX Improvements to the* Zulip Terminal *interface

Supervisor: *Prof. Dibyendu Chakarabarty (PI, ISRO Aditya-L1 Mission)* Spring 2023

- Wrote 12 Pull-requests, introduced mechanisms for live read-receipts, last-seen counters, and assisted in documentation. Resigned after mid-evaluation to focus on research.

Granular.AI | Machine Learning Consultant [Summer 2023]

Zero-Shot Learning for Satellite image Segmentation and Captioning

Supervisor: *Sagar Verma (CTO, Granular.AI), Maria Vakalopoulou (Université Paris-Saclay)*

- Worked on the generalizability and out-of-the-box training capability of foundational models on limited-exposure datasets like geospatial information for downstream tasks.

Snapper AR | Software Developer [2022-23]

Wireless video-streaming pipeline for ML-based AR filters from Nvidia Jetson

Supervisor: *Imtiyaaz Ansari (Founder, Snapper AR)* Winter 2022-23

- Integrated `GStreamer`, Nvidia `Deepstream`, `V4L2` into an Nvidia Jetson, established connection between Jetson and Arducam IMX477 and built a web-app for streaming CUDA-accelerated model inferences to a connected smartphone with minimal lag. My tool was presented in an investment pitch that resulted in 2 Million INR in funding for Snapper.

Defense Research and Development Organisation, India | Research Intern [2023]

Reinforcement Learning Policies for Drone Swarms

Supervisor: *Dr. Fannikiran Maddukuri* Spring 2022, Summer 2022

- Built a Reinforcement Learning scheme for a `pybullet` environment with simulated physics for automated learning of drone movement on `ROS`; designed effective policies via rapid experimentation with an analogous simulation designed on on Unity with `MLAgents` library.

Physical Research Laboratory, ISRO, India | Research and Development Intern [2023]

Early Warning and Forecasting systems for inbound Geomagnetic Storms

Supervisor: *Prof. Dibyendu Chakarabarty (PI, ISRO Aditya-L1 Mission)* Spring 2023

- Implemented a Transformer pipeline with Bayesian Inference for data from India's maiden solar probe, the *Aditya-L1* to predict anomalous geomagnetic storms with uncertainty quantification. **Submitted to** Advances in Space Research **journal**, **and under deployment at** ISRO.

Selected Open-Source Projects

- `PineappleEMU` | **Accurate Emulator for the Nintendo Entertainment System**
  Mentor: *Prof. Sameer G. Kulkarni*, IIT Gandhinagar *Course Project*
  Accurate emulation for all ∼70 opcodes for the MOS 6502 CPU, and semi-accurate emulation of the the NTSC 2C02 PPU through a pixel-blitting engine, written in C++ for x64 Intel hardware.
- `GossipLang` | **Pedagogical Multi-Paradigm Programming Language.**
  Mentor: *Prof. Balagopal Komarath*, IIT Gandhinagar *Course Project*
  Object-oriented language with first-class functions, with both a tree-walk and bytecode interpreter, with pedagogical visualizations for abstract syntax trees and variable resolution pass written in `python`; to be used in future offerings of the Compilers course as a teaching tool.
- `WimpyML` | **Machine Learning library written from scratch.**
  `WimpyML` is a library for machine learning algorithms written from scratch in `numpy` and `JAX`, including popular regression techniques, sampling algorithms, neural networks, and optimizers.
- `GPZoo` | **Gaussian Process paper implementations, written from scratch.**
  `GPZoo` consists of major advances in classical Gaussian Processes, written from scratch in `torch` and `numpy`, including Neural Kernel Networks, Sparse Variational GPs, and more.
- `InsIIT` | **Official IIT Gandhinagar Institute App, used by 2000+ students daily.**
  Flutter-based app for dynamically tracking what is happening on the IIT Gandhinagar campus, used by over 2000 students and members of the IIT Gandhinagar community on a daily basis.

I have also contributed extensively to `zulip-terminal` as part of Google Summer of Code 2022, over May to September, 2022. I like making libraries, and my Github has a number of cool ones.

Leadership Positions & Social Service

- Technical Secretary, Student Council, IIT Gandhinagar
  Elected by a electoral pool of 2000+ students, and managed a team of 50+ students in 7 technical clubs, formulated induction assignments for clubs, conducted 16 workshops, 10+ events and guidance sessions, which resulted in 4 student-only research papers and the historical best performance of our students in various international contests.
- Core Committee Member, Blithchron, Cultural Fest, IIT Gandhinagar
  Head of Sponsorships in the Cultural Fest committee of IIT Gandhinagar, 1 of 4 student leaders leading a team of 140+ to conduct Gujarat's first fully-offline cultural fest after the COVID-19 pandemic. Garnered ≈ 1 Million INR in sponsorships from 20+ companies.
- Industrial Relations Coordinator, IIT Gandhinagar
  Oversaw a 400% increase in the number of industrial internships and collaborations in the 2022-23 cycle, with a select set of 15 students, with 500+ companies reached and 50+ collaborations. Coordinated and managed the first continuous documentation effort for all team activities.
- General Member, Student Academic Council, IIT Gandhinagar
  Oversaw multiple student-centric activities, research seminars, and chaired outreach activities for the student community. and taught IIT Gandhinagar's first student-run credited course, on Mathematics for Machine Learning, with 170+ registered students.
- Theatre Artiste, Tripura Puppet Theatre, Tripura
  Member of a 10 pax team involved in contemporary puppet research, design and performance. Performed at 10+ national and international venues, notably Days of India in Germany, 2012 at Hamburg, Dresden and Berlin, celebrating 75 years of democratic relations between India and Germany. Awarded a national scholarship for my involvement in this artform.