Ayush Raj

BS-MS dual degree programme, Indian Institute of Science Education and Research (IISER) Pune

raj.ayush@students.iiserpune.ac.in | github.com/Artamta | linkedin.com/in/ayush-raj-7608a424a

Profile

I am a 5th year **Student researcher** passionate about addressing real-world problems using data-driven techniques. My expertise lies in **Machine learning, Deep learning using HPC**. I focus on designing and implementing Parallel, scalable, reproducible pipelines for complex multidisciplinary research.

Research Experience

MS-Thesis Researcher, SCDLDS, Ashoka University
Master's Thesis: **AI Subseasonal-to-Seasonal Weather Forecasting for Precipitation**
Supervisor: Dr. Sandeep Juneja

- Developing a novel **AI Subseasonal-to-Seasonal** weather forecasting model specilized for Indian Region.
- Building **IndiaS2S-Bench**, a unified s2s model comparison over India and Homogenous regions
- Evaluating AI forecasts against operational dynamical models using **ACC, RMSE, Bias, CRPS, and CRPSS**.
- Timeline: **May 2026-Present**

Research Project Student — Transformer-based Weather Forecasting, IISER Pune
Supervisor: Dr. Bedartha Goswami

- Implemented a transformer-based weather forecasting model for spatio-temporal climate prediction using the **FuXi architecture** in PyTorch.
- Pretrained the model on **40 years** of global atmospheric reanalysis data (**1.5**° resolution) comprising **20 meteorological variables**, including surface variables and upper-air fields at **850, 500, and 250 hPa**.
- Trained and evaluated multiple model configurations, performing hyperparameter optimization and ablation studies to improve forecasting accuracy and computational efficiency.
- Performed distributed multi-GPU training on **NVIDIA A30 and A100** clusters using **FSDP**, Hugging Face Accelerate, mixed-precision training, and gradient checkpointing.
- Achieved an **Anomaly Correlation Coefficient (ACC) > 0.5** for weather forecasts up to **5 days** ahead.
- Repository: github.com/Artamta/Fuxi-Weather-Prediction

Summer Research Intern — Quantitative MRI, Centre for Biomedical Engineering, IIT Delhi (Remote)
Supervisor: Dr. Amit Mehndiratta (IIT Delhi & AIIMS Delhi)

- Developed voxel-wise **IVIM-DKI** models for quantitative diffusion MRI analysis using real patient datasets across multiple anatomical regions.
- Implemented nonlinear curve fitting, parameter estimation, and mathematical modeling for diffusion MRI using **SciPy**, **lmfit**, and Python.
- Designed end-to-end preprocessing and analysis pipelines for DICOM/NIfTI datasets using **NiBabel**, **SimpleITK**, and **MRICron**.
- Evaluated hybrid **IVIM-DKI with Total Variation regularization**, demonstrating a **30% improvement** in parameter-map stability under low-SNR conditions through synthetic simulations and clinical studies.
- Performed quantitative evaluation using **RMSE**, **Bias**, and **AIC** on liver, lymphoma, prostate, and rectal MRI datasets.

Research Project Student — Spatio-Temporal Climate Data Analysis, IISER Pune
Supervisor: Dr. Bedartha Goswami

- Conducted spatio-temporal analysis of extreme heat events across the Indian subcontinent using **40 years (1979–2018)** of ERA5 2-m temperature reanalysis data at **0.25**° spatial resolution.
- Applied unsupervised machine learning techniques, including **K-Means** and **UPGMA clustering**, to identify and classify seasonal heatwave patterns.
- Utilized the **DeepGraph** framework to model complex spatial and temporal relationships in large-scale climate datasets.
- Investigated the impact of extreme heat events on ecosystem and vegetation dynamics through data-driven climate analysis.

- Developed HPC-based workflows for large-scale climate data processing, geospatial analysis, and environmental modelling using **Python**, **Xarray**, and **DeepGraph**.
- Repository: github.com/Artamta/Spatio-Temporal_Heatwave_Analysis

Education

Indian Institute of Science Education and Research (IISER), Pune
BS-MS dual degree programme

Jawahar Navodaya Vidyalaya, Pithoragarh
10th: 93.6%, 12th: 93.2% | CBSE Merit Certificate (Science)

Relevant Coursework

- **Mathematics:** Probability and Statistics, Calculus I & II, Linear Algebra, Markov Chains, Numerical Computation,Graph-Theory
- **Computer Science:** Systems and Databases, Systems and Implementation of Algorithms
- **Data Science & AI:** Data Analysis,Statistical Learning and Data Science, Deep Learning, Bioinformatics,Image and Video Processing with Deep Learning

Selected Projects

Curriculum Graph Studio:

- Developed a graph-based curriculum planning dashboard for IISER Pune,
- Repository: github.com/Artamta/project/cg

Facial Emotion Recognition System (Course Project: Image & Video Processing with Deep Learning)

- Designed and implemented a deep learning pipeline for emotion recognition using the **CK+ dataset** (7 classes: Angry, Disgust, Fear, Happy, Sad, Surprise, Contempt).
- Built on a **VGG19 backbone**, adapted for grayscale inputs, with dropout and weight decay for regularization.
- Implemented a **5-fold cross-validation pipeline** in PyTorch, achieving an average test accuracy of **73.9%** and a best fold accuracy of **79.8%**.
- Repository: github.com/Artamta/project_ayush_raj

Breast Cancer Detection Model (Practice Project)

- Built and evaluated a logistic-regression classifier on clinical features (97% accuracy on held-out test); deployed a Flask demo for interactive use.
- Repository: github.com/Artamta/Brest_Cancer_Detection_Model

Extracurricular Leadership

Hackathons, Study Groups Open Source: Active contributor to coding initiatives, collaborative projects, and open-source research software. **NVIDIA Earth-2:** Contributed the FuXi-S2S prognostic model to the NVIDIA Earth-2 ecosystem; actively contributing to the S2S forecasting initiative in collaboration with IMD. *Project:* s2s.scdlds.com Represented IISER Pune in inter-college cricket tournaments; Captian Cricket Team.

Technical Skills

Programming: Python (**Proficient — 2.5+ years**), C++, C, SQL, Bash, SSH
Frameworks / ML: PyTorch (**Proficient**), Scikit-learn, Transformers, Hugging-face, LLMs, RAG, LangChain
Data / Geo / Imaging: NumPy, Pandas, Xarray, GeoPandas, NiBabel, SimpleITK
High-performance computing: Slurm (job scheduling, cluster experiments), HPC workflows
Tools: Git, Linux, Matplotlib, Seaborn, Flask, Heroku

Languages

Hindi (Native), **English (Fluent)**, **Kumauni (Fluent)**,**German (Beginner)**