SARTHAK ARORA

sarthakvarora@gmail.com | P: +91-8081880257 | sarthakvarora.github.io

EDUCATION

PLAKSHA UNIVERSITY — Chandigarh, India
Post Graduate Diploma — 2021 – 2022
Major in Computer Science — CGPA: 3.88/4.0; Department Rank 2
Relevant Coursework: Machine Learning; Computer Vision; Reinforcement Learning; GraphML

RAMJAS COLLEGE, DELHI UNIVERSITY — New Delhi, India
Bachelor of Science — 2015 – 2018
Major in Statistics; Minors in Economics — CGPA: 7.176/10.0; First Division
Relevant Coursework: Actuarial Sciences; Time-Series Analysis; Statistical Modelling; Hypothesis Testing

PUBLICATIONS

[Application of Analytical Hierarchical Process and its Variants on Remote Sensing Datasets](#) - ArXiV preprint
Sarthak Arora, Michael Warner, Ariel Chamberlain, James C. Smoot, Nikhil Raj Deep, Claire Gorman, Anthony Acciavatti

[Queer In AI: A Case Study in Community-Led Participatory AI](#) - FAccT'23 ***Won best paper** 🏆
Sarthak Arora, Organizers of Queer in AI

[Was that Sarcasm?: A Literature Survey on Sarcasm Detection](#) - Presented at NAACL'24
Sarthak Arora, Harleen Kaur Bagga, Jasmine Bernard, Sahil Shaheen

RESEARCH EXPERIENCE

COLLABORATIVE EARTH — Remote
ML Engineer — Jan 2023 – Present
- Using **Computer Vision** and GEE to identify and cluster "Naalas" for the Ganges river from Satellite Images.
- Used GIS Tools and datasets to create a comprehensive **site suitability** map for Coastal Forests Wetlands Lab.
- Conducted **fieldwork** around the Ganges River which involved drone surveys and meeting stakeholders.
- Leading **fundraising** efforts for the Ganges Lab, finalists for the [Frontier](#) CDR Prepurchase 2024.

WUUII INC. — Remote
ML Engineer — Jun 2024 – Present
- Utilizing a double-encoder double-decoder **Vision Transformer** model to create a shared latent space between streetview and aerial data to identify features across modalities with enhanced geolocation.
- Fine-tuned Segment Anything on streetview imagery and translated the identified features onto an aerial map to integrate polygon features and depth information for improved Wildfire Risk Assessment.
- Representing the spread of wildfires using a **Graph Attention Network** model to understand factors that facilitate fire propagation and inform preventive strategies.

UNIVERSITY OF CALIFORNIA, BERKELEY — Remote
Research Scientist — Sep 2022 – Present
- Conducting a time-series **Anomaly Detection** of Powerplant Emissions to find areas of policy improvement.
- Analyzed Quantitative and Qualitative Regressors for a **Wildfire Survey** conducted by [MWPA](#).

UNIVERSITY OF CALIFORNIA, BERKELEY — Berkeley, CA
Research Scholar — Apr 2022 – Jul 2022
- Conducted Research to develop a wildfire **risk assessment** tool under the mentorship of Dr. Thomas Azwell.
- Worked on **Remote Sensing** datasets for the state of California to predict wildfires with an accuracy of **82%**.

AI FOR GOOD FOUNDATION   Remote
ML Engineer   Feb 2022 – Feb 2024
- Worked on the Climate SDG project, backed by the UN Development Program and was presented at COP28.
- Created a Graph based Visualisation for Climate Entities using Named Entity Recognition and Neo4J.
- Cleaned Scraped data using LDA for topic modeling, leading to a Recall value of **83.5%** in noise detection.

QUEER IN AI   Remote
Organizer   Jan 2022 – Present
- **Program Chair** for the Queer in AI Workshop at **NeurIPS** for the years 2022-2024.
- Working on the [AI Harms and Proposed Policy Interventions](#) to identify areas for AI policy improvement.
- Conducting and publishing research to raise awareness of queer issues in the general AI community.
- Working on the Graduate Application Aid program, aimed at supporting queer applicants through funding and mentorship, led to scholarships worth **$130,000** awarded to **130+** applicants across **30** countries.

INDUSTRY EXPERIENCE

EIRENE AI   Remote
ML Engineer   Jul - Sep 2024
- Fine-tuned **RT-DETR** to detect fire and smoke in real-time, to be deployed on-edge on CCTV cameras.

PREDLI AB   Remote
ML Engineer   Sep 2023 – Feb 2024, May - June 2024
- Improved LLM Performance on a Tweet generator using a **RLHF** model based on Tweet engagement by 10pp.
- Created a Machine Learning Model trained on Geophysical Layers such as Bedrock, Magnetism, Resistivity, Gravity and Satellite Imagery for Mineral **Site Prediction** in Sweden with an accuracy of **89%**.

INDIFI TECHNOLOGIES   Gurgaon, India
Data Scientist   Aug 2022 – Jul 2023
- Created a model which predicts Default Risk with **96%** precision by analyzing CIBIL Trend using **LSTM**.
- Used **Fuzzy Logic** to match addresses for verification, leading to a reduction in manual checks by ~4 hours.

PROJECTS

- Created a weather prediction tool using FourCastNet and GPT for the **OpenAI** Climate Hackathon.
- Created a real-time face and facial-features detector on live videos using **HAAR** and **CNN.**
- Created a Machine Learning model for Drug Discovery using **GANs**.

TALKS AND COURSES

- Delivered a talk on "South Asian Feminism/s in the 21st Century" at **ICML 2022**.
- Completed the Hilary 2024 School of Climate Change course offered by **Oxford Climate Society**.
- Completed the **Climate Change AI** Summer School 2024.
- Cleared Actuarial papers **CT1** (Financial Mathematics), **CT3** (Probability and Mathematical Statistics), **CT4** (Models), **CT5** (Contingencies), **CT6** (Statistical Methods) from the Institute and Faculty of Actuaries (IFoA).

TEACHING

- Taught subjects such as ML, Research, and Leadership to students based out of Africa with **Belong.**
- **Teaching Assistant** for the Economics Courses at Ramjas College, Delhi University.

ACHIEVEMENTS

- Silver Medalist for academic excellence at Plaksha University, TLP Batch of 2022.
- Winning Team for the Plaksha Entrepreneurial Challenge Lab where we pitched a Mental Health startup.
- Won the Mercy Corps Hackathon for predicting commodity prices for Disaster relief in Africa using LSTMs.