

# Swapnil Mukherjee

 [linkedin.com/in/swapnil-mukherjee](https://www.linkedin.com/in/swapnil-mukherjee)

 [github.com/SwapnilMukherjee](https://github.com/SwapnilMukherjee)

 [swapnil.mukherjee.ug25@ashoka.edu.in](mailto:swapnil.mukherjee.ug25@ashoka.edu.in)

 [swapnailmukherjee.github.io](https://github.com/swapnailmukherjee)

## EDUCATION

---

- **Ashoka University**  
BSc. (Hons.) in Computer Science (GPA: 3.71/4)

Sonepat, Haryana

2022 - 2026

## RESEARCH INTERESTS

---

Multimodal Learning, Reasoning in AI, Neurosymbolic Methods, World Models, Commonsense Understanding

## PUBLICATIONS/PRE-PRINTS

---

1. Aritra Dutta, **Swapnil Mukherjee**, Deepanway Ghosal, and Somak Aditya. (2025). *NLKI: A Lightweight Natural Language Knowledge Integration Framework for Improving Small VLMs in Commonsense VQA Tasks*. In Findings of the Association for Computational Linguistics: EMNLP 2025, pages 10543–10563.
2. Debojyoti Chowdhury, Bhavesh Neekhra, Shreyansh Priyadarshi, **Swapnil Mukherjee**, Debashruti Maity, Debayan Gupta, and Shubhasis Haldar. (2023). *Next-Gen Profiling of Tumor-resident Stem Cells using Machine Learning*. bioRxiv.
3. **Swapnil Mukherjee** and Sujit Das. (2021). *Application of Transformer-Based Language Models to Detect Hate Speech in Social Media*. In Journal of Computational and Cognitive Engineering, 2(4), 278-286.

## WORK EXPERIENCE

---

- **Precog Lab, IIIT Hyderabad**

*Research Intern*

Hyderabad, India

June 2025 - Present

- ◊ Working under Prof. Ponnurangam Kumaraguru (Precog), in collaboration with Tanuja Ganu (Microsoft Research India) to explore modality bias and preference in Multimodal ('omni') LLMs that can process inputs and produce outputs across text, image, audio, and video.
- ◊ Our preliminary findings reveal that these models exhibit uneven capabilities across modalities, performing non-trivially better in some modalities than others for the same tasks. I am leading the development of a benchmark for omni-modal composition and counterfactual reasoning abilities in these models.

- **SarvamAI**

*ML Engineer Intern*

Bangalore, India (Remote)

June - August 2024

- ◊ Independently developed the first prototype of Sarvam's multilingual document text recognition system for 10+ Indic languages. Set up data curation pipelines, large-scale multi-node training, and experimented with different novel encoder-decoder architectures and modality fusion strategies.
- ◊ Contributed to Sarvam's Parsing API by benchmarking, evaluating, and generating training data for the underlying models.

- **Koita Centre for Digital Health - Ashoka (KCDH-A)**

*Research Intern*

Sonepat, India

March - October 2024

- ◊ Worked with Prof. Debayan Gupta and Prof. Steven Clipman of Johns Hopkins Medicine Institute on machine learning-based methods for enhanced forecasting of critical antiretroviral therapy (ART) drugs for the National AIDS Control Organization, Ministry of Health.
- ◊ Set up the data acquisition and processing pipeline and experimented with various classical and neural network time-series models. Developed the first few iterations of the prediction model, and presented our initial results in a presentation to senior executives and clinicians at NACO.
- ◊ Our method is being adopted by NACO for country-wide forecasting of 17+ HIV drugs which will optimise drug availability and impact 2.5M patients living with HIV in India.

## RESEARCH EXPERIENCE

---

- **Undergraduate Thesis on Neurosymbolic Reasoning for VQA:** I am extending the framework proposed in NELLIE and TV-TREES family of work to enable grounded and verifiable reasoning for Visual Question Answering (VQA), supervised by Prof. Debayan Gupta and Prof. Somak Aditya. I am primarily working to (*i*) expand the scope of this framework to open-domain commonsense-based image-QA, (*ii*) contribute a novel ‘Visual Grounding Module’ to create proof-trees grounded directly in image content (*iii*) use small, local models instead of large closed-source models in the engine. (*In Progress*)
- **Commonsense Reasoning in Vision-Language Models:** Collaborated with Prof. Somak Aditya (IIT Kharagpur) on the NLKI framework, which enhances small Vision-Language Models (sVLMs) for commonsense VQA tasks. By integrating externally retrieved knowledge and LLM-generated explanations, our method improved the performance of sVLMs by up to 15%, enabling them to match or outperform larger generative models.  
[Paper] [Code] [Webpage]
- **Precision Profiling of Tumor Stem Cells Using Machine Learning-Based Digital Cytometry:** Worked with Prof. Debayan Gupta and Prof Subhashis Halder’s lab to develop an ML-based method for classifying and quantifying gene expression data from human stem cells, enabling highly accurate profiling of cancerous tumors to aid diagnosis and treatment. Responsible for data collation and preparation, running all ML experiments, and developing the core codebase for a software package that integrates this functionality into a user-friendly tool. Patent app. no. 202511064233. (*Under review*)  
[Pre-print]
- **Coreset Selection for Image Datasets:** Under the supervision of Prof. Raghavendra Singh, I developed a novel zero-shot method for coreset selection (on image datasets) using a simple graph-based approach for dataset distillation. The highlight of the method is that can identify important examples in a given dataset without requiring any training on the target dataset. It is comparable to current SOTA methods while being much more computationally simple and efficient. (*Sep ’23*)
- **Miscellaneous:**
  - Created an **Agentic AI system for no-code AutoML** that autonomously runs end-to-end data analysis or ML-modelling pipeline on any dataset without user supervision. [Code]
  - Worked on a Causal Reinforcement Learning project combining Deep Q-Learning and Bayesian Inference, which taught the agent **explicit mechanics of causality** in exploration through a bayesian belief matrix and drastically improved generalization to completely unseen environments. [Code]
  - Developed a minimal prototype of a system **VeriDoc** for verifiable integrity in public document archives. [Code]
  - Presented our poster on machine learning-based enhanced forecasting of antiretroviral therapy demand in India at the Johns Hopkins GKII Meet (2024) **amongst 50+ groups sponsored by the GKII Breakthrough grant** in India.
  - Core contributor on a poster titled *FedTree: Federated Learning on Tabular Medical Data* accepted and presented at the 4th Indian Cancer Genome Atlas (ICGA) Conference 2023. [Poster]

## SKILLS AND TOOLS

---

- **Programming Languages:** Python, C, Java, SQL, HTML, CSS
- **Language Proficiency:** English (full professional), Bengali (native), Hindi (full professional)
- **Tools:** PyTorch, Transformers/HF, Pytorch Lightning, LlamaFactory, TransformerLens, Ollama, vLLM, CrewAI, Keras, Numpy, SKLearn , SciPy, Pandas, Matplotlib, Gradio, Git, Docker, LATEX

## HONORS AND AWARDS

---

- Science Academies’ Summer Research Fellowship Programme 2025 Fellow
- Dean’s List (for academic excellence in a semester; **thrice**).
- College Board India Scholar 2022.
- CCIR Summer TechCamp Distinguished Scholar Award for exemplary performance in the programme (Aug ’21).