Swapnanil Mukherjee

[in] linkedin.com/in/swapnanil-mukherjee

[✉] swapnanil.mukherjee_ug25[at]ashoka.edu.in

[O] github.com/SwapnanilMukherjee

[🌐] swapnanilmukherjee.github.io

EDUCATION

Ashoka University — Sonepat, Haryana
Bachelor of Science, Computer Science (3.66/4) — 2022 - 2026

RESEARCH INTERESTS

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

PUBLICATIONS

1. Aritra Dutta, **Swapnanil 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, **Swapnanil Mukherjee**, Debashruti Maity, Debayan Gupta, and Shubhasis Haldar. (2023). *Next-Gen Profiling of Tumor-resident Stem Cells using Machine Learning*. bioRxiv.

3. **Swapnanil 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 — Hyderabad, India
Research Intern — 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. Ongoing work to create a benchmark that can quantify this limitation and provide a true measure of multimodal fusion and understanding in these models.

SarvamAI — Bangalore, India (Remote)
ML Engineer Intern — 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) — Sonepat, India
Research Intern — March - October 2024

◇ Working with Prof. Debayan Gupta in collaboration with 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 cleaning and processing pipeline and experimented with different models. Developed the first few iterations of the forecasting tool. Presented our initial results in a presentation to senior executives at NACO.

◇ Our method is being adopted by NACO for country-wide forecasting of 17+ HIV drugs which will optimise drug availability and positively impact lakhs of 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 more expleci leverage visual information (*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](https://www.shubhasis-haldar-cmt.com/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**: Came up with 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**:
  - 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. [**Webpage**]
  - Led a research project for NeuroMatch Academy 2022 to explore whether the behaviours of the different visual cortices of a human brain upon exposure to a visual stimulus could be imitated by the layers of a convolutional network. [**Code**]
  - Developed the prototype of an emotion recognition module during my internship at BrainEnTech Neuroscience. [**Code**]

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, vLLM, 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 exeplary performance in the programme (Aug '21).