Mayukh Deb

mayukhmainak2000@gmail.com mayukhdeb.github.io mayukhdeb mayukhdeb

Science

TopoNets (ICLR 2025 Spotlight → toponets.github.io) Jan 2025

Mayukh Deb, Mainak Deb, N. Apurva Ratan Murty
Inducing topographic structure in Vision and Language models. Yielded brain-like functional organization, parameter efficiency and lower dimensionality.

AtMan (NeurIPS 2023 + featured in Scientific American) Jan 2023

Mayukh Deb*, Björn Deiseroth*, Samuel Weinbach* et al. (* = equal contribution)
Causally trace and explain LLM outputs without gradients. Works on anything with attention. Foundation behind Aleph-Alpha's Explain functionality

DORA (ICLR 2023 Trustworthy ML workshop + TMLR) June 2022

Kiril Bykov, Mayukh Deb, Klaus Robert Müller et al.
Clustering neurons and detecting spurious outlier features with feature-vis.

Education

Georgia Institute of Technology *Aug 2024 – Present*
PhD Student - Cognition and Brain Science - Murtylab

- Mentored by Dr. N. Apurva Ratan Murty
- Developing brain-inspired algorithms to improve Language and Vision models (see recent work: toponets)
- Leading the engineering efforts behind training and hosting of state-of-the-art models trained on fMRI data

Experience

Research Engineer @ Aleph-Alpha *Nov 2021 – May 2023*

- Led their Explainable AI project and built AtMan
- AtMan was the foundation behind Aleph-Alpha's "explain" API for LLMs
- Also worked on building multimodal search-engines.

Research Intern @ MIT Brain + Cognitive Sciences *May 2023 - Dec 2023*

Worked with Dr. Nancy Kanwisher's lab on 2 projects:

- Inducing brain-like topographic structure in transformers (eventually led to toponets)
- Training data-constrained vision models on fMRI data

Research Engineer @ Eden.Art *Dec 2023 – Aug 2024*

- Worked with stuff like Textual Inversion, IP-Adapters, ZipLoRA etc.
- Built highly flexible pipelines to fine-tune diffusion models (SDXL, SD3) quickly on user data

Intern @ RunwayML *Jan 2021 – Feb 2021*

- Implemented, optimized (1.4x speedup) and dockerized pipelines for optical-flow (RAFT) and video frame interpolation (RIFE) models to be used in Runway's video editing tool.

Google Summer of Code @ INCF *May 2020 – Aug 2020*

- Worked with OpenWorm to train models to extract metadata from microscopic videos/images embryos
- Also mentored two contributors in GSoC 2023.

Open Source

TopoLoss

- Induce topographic structure in pytorch models during training with this loss function
- Works on both Linear and Conv layers
- Core codebase behind toponets
- Also released some pre-trained vision and language models

torch-dreams

- A highly flexible framework to do feature visualization on pytorch models
- **Over 52k downloads** on PyPI

Eden

- Single python decorator to convert a python function into a hosted endpoint with queuing (celery)
- Surprisingly scalable across instances with kubernetes
- Foundational pet-project which eventually led to eden.art
- 8.4k downloads on PyPI

DevoLearn

- Trained models to segment embryo data from microscope
- Outcome of Google Summer of Code, 2020 and then taken forward by other students in the next years

More projects can be found on my github profile: github.com/mayukhdeb

Technologies

Languages: Python and a little bit of CUDA – I just learn whatever is required

Frameworks: PyTorch, NumPy, Pandas