Mayukh Deb

mayukhdeb.github.io | github.com/mayukhdeb

PUBLICATIONS

AtMan: Understanding Transformer Predictions Through Attention Manipulation

January 2023

Mayukh Deb*, Björn Deiseroth*, Samuel Weinbach*, Manuel Brack, Patrick Schramowski, Kristian Kersting (* = Equal Contribution)

Accepted at NeurIPS 2023 + Featured in Scientific American

DORA: Exploring outlier representations in Deep Neural Networks

June 2022

Kirill Bykov, **Mayukh Deb**, Dennis Grinwald, Klaus Robert Muller, Marina MC Höhne
ICLR 2023 Trustworthy ML Workshop + Transactions in Machine Learning Research

EXPERIENCE

PhD Student - MurtyLab @ Georgia Tech

Aug 2024 - Present

Research Engineer @ eden.art

December 2023- Aug 2024 (Remote)

Working with stuff like Textual Inversion, IP-Adapters, ZipLoRA and DPO and Diffusion Attentive Attribution Maps (DAAM). Also fine-tuning state models like SDXL, SD3.

Research Intern @ MIT

May 2023 - December 2023 (Remote)

Worked with **Dr. Nancy Kanwisher's Lab** (primarily with Dr. N Apurva Ratan Murty) on 2 projects:

1. Training deep neural networks on fMRI data
2. Inducing brain-like topographic properties in transformers

Research Engineer @ Aleph-Alpha

May 2022 - May 2023 (interned from November 2021 to April 2022) (Remote)

Led the efforts to reverse engineer Large (Multimodal) Language Models for Interpretability. Our efforts eventually led to the release of AtMan as a product offering for Aleph-Alpha's enterprise customers.

Google Summer of Code 2020 @ INCF

May 2020 - August 2020 (Remote)

Collaborated with OpenWorm to train deep neural networks to extract useful metadata from microscopic videos/images of the c. elegans embryo. Also mentored 2 contributors in 2023

Intern @ RunwayML

January 2021 - February 2021 (Remote)

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.

SKILLS

PyTorch
NumPy
Figuring things out

LANGUAGES

English
Bengali
Hindi

OPEN SOURCE/PET PROJECTS

Torch-Dreams

September 2020 - Present

Torch-Dreams is a tool used to reverse engineer vision models in PyTorch using feature visualization. It is actively used by research groups like UMI Lab@TU Berlin and KanLab@MIT

Devolearn

August 2020 - Present

DevoLearn is a python library which helps accelerate data driven research on embryos with Pre-Trained deep learning models used for feature extraction directly from microscopic videos and images. It's now an open-source active project under INCF with Google Summer of Code

Eden

July 2021 - Present

Eden helps to deploy python functions as a hosted endpoint with support for multiple GPUs with minimal changes to the existing code. It used to power the AI art pipelines behind abraham.ai and was the foundational pet-project which eventually led to eden.art

WRITING

Learning to Make the Right Mistakes - a Brief Comparison Between Human Perception and Multimodal LMs

July 2021

This article was written for TheGradient. It can be found here.

HACKATHONS

3rd Prize: MLOps for Good Hackathon organized by Microsoft, Iguazio

July 2021

Built deepfake-shield, a web-app to detect faces from within an image and determine if they're deep-fakes or not.

OTHER ODD PROJECTS

Differentiable-morphogenesis

March 2021

Implemented differentiable self organizing systems using neural cellular automata on PyTorch.

Patrick

December 2020

Tiny neural net library written from scratch with CuPy.

Gradient Games

August 2020

Wrote simple neural networks on numpy from scratch along with their backpropagation algorithms and visualized the weights as they trained.

Deep chicken terminator

February 2020

Tracked various animals using deep-learning from the live feed of the player's perspective in the game Minecraft.

Deep chicken saviour

March 2020

Generated "adversarial textures" within the game which were misclassified as animals by the model from deep chicken terminator.

Mind the bend

December 2019 - August 2020

Made a car steer by itself in a 3D racing game using an ensemble model and some automation with PyAutoGUI. The model was trained on custom data which comprised of images from two perspectives of the road for better performance

EDUCATION

B. Tech, Electrical and Computer Engineering, Amrita Vishwa Vidyapeetham

2019-2023

CGPA: 7.95

VOLUNTEER WORK + EXTRACURRICULAR ACTIVITIES

Black Belt in Kyokushin Karate

I've been a part of the martial arts class in my school and participated in multiple tournaments in Kolkata, India under the mentorship of Sensei Bejoy Dhara, branch chief and country representative for IKO Matsushima.

Co-ordinated charity events to support villages

Helped put together events where students donated basic medical supplies and other utilities which were to be given to the poor and needy children in villages near Kolkata, India. This programme was a part of Mata Amritanandamayi devi's (Popularly known as "Amma") efforts to help the poor and the needy throughout the world.

Amala Bharatam

Volunteered on clean-up drives in Kolkata as a part of the "Amala Bharatam" programme under Mata Amritanandamayi Math