Ayudh Saxena

ayudhs@cs.cmu.edu | ayudhsaxena | ayudhsaxena.github.io

Research Interests

Representation learning, Multimodal ML, Mechanistic Interpretability, Social AI

Education

Carnegie Mellon University, Pittsburgh
M.S. in Machine Learning — 2024 – 2025
Courses: Representation Learning, Adv. Intro to ML, Prob. & Statistics — GPA: 4.0

Indian Institute of Technology (IIT), Kharagpur
Bachelor of Technology (Hons.) in Computer Science and Engineering — 2018 – 2022
Advisor: Prof. Sandip Chakraborty — GPA: 9.43/10.0

Patents & Publications

[1] vBeats:A Framework for Converting Gestures into Bass Strokes from Indian Tabla using Smartphones
A. Saxena, S. Chatterjee, S. Chakraborty - **Best Poster Award**
Accepted at COMSNETS: 16th International Conference on Communication Systems & Networks 2024 [PDF]

[2] Videos2Doc: Generating Documents from a Collection of Procedural Videos
T. Shukla, A. Bhattacharyya, **A. Saxena**, J. K. Karnuthala, A. Bohra, B. P. R. Guda, A. Sancheti, N. Chhaya
In ACM IUI '22 Companion: 27th International Conference on Intelligent User Interfaces 2022 [PDF]

[3] Automated Digital Document Generation from Digital Videos
T. Shukla, A. Bhattacharyya, **A. Saxena**, J. K. Karnuthala, A. Bohra, B. P. R. Guda, A. Sancheti, N. Chhaya
US Patent No. 11,783,584 · Issued 10th October, 2023 [PDF]

Research/Experience

Adobe Inc. - Software Development Engineer II, Photoshop Express — July '22 - Aug '24
Domain: Software Development, Operating Systems, Machine Learning

- Designed and developed a full-fledged video editing segment within **Photoshop Express** as a founding member of the Video Team, utilizing **SwiftUI** framework to create a seamless user experience.
- Implemented a novel solution to generate contextually appropriate, stylistically consistent meme captions for images utilizing cloud-based image-to-text models followed by GPT using **Chain of Thought** prompting.
- Recipient of the **Adobe Spot** and **Adobe Star** award for excellent engineering work done for the year 2022-23.

Undergraduate Researcher - Bachelor Thesis Project — Aug '21 - May '22
Research Areas: Representation Learning, Human-Computer Interaction — *Advisor:* Prof. Sandip Chakraborty

- Developed a novel approach for converting Indian Tabla gestures into bass strokes using smartphone sensors. Tackled the problem of **lossless conversion between two modalities** having vastly differing sampling rates.
- Proposed **vBeats [1]**, a multi-modal framework that uses autoencoders to learn meaningful latent spaces for inertial and audio signatures, followed by LSTMs to effectively map representations using seq2seq learning.
- Curated a dataset of 1.38 hours of synchronized IMU-Audio signals with 4 professional tabla players.
- Nominated for **best departmental thesis**.

Adobe Research - Research Intern, Big Data Experience Lab — May '21 - Aug '21
Research Areas: Multimodal Learning, Intelligent User Interfaces — *Team:* Big Data Experience Lab

- Developed **Videos2Doc [2]**, a machine learning-based framework for generating documents from collections of procedural videos, tackling the task of extracting information from multimodal content.
- Implemented a complex pipeline that extracts the best video-frames by constructing an action graph that is then fed into an encoder-decoder-decoder model along with textual ingredients data to generate instructions.

Technical Skills

Languages: Swift, Python, C++, Objective-C, C, Java, Scala, SQL, MIPS, Verilog HDL, HTML

Libraries: PyTorch, Tensorflow, pandas, SciPy, Numpy, SwiftUI, Spark

PROJECTS

Extracting Monosemantic Multimodal Features using k-SAE [Report] Fall 2024
Representation Learning *Guide:* Prof. Pradeep Ravikumar
- Extracted interpretable multimodal features from vision, audio, and text embeddings using a k-Sparse Autoencoder

SemEval-2022 Task 11: MultiCoNER [Report] Spring 2022
Natural Language Processing *Guide:* Prof. Pawan Goyal
- Employed paraphrasing while preserving the original sentence's meaning and ensuring entity preservation, as a data augmentation technique to improve the performance from an F1 score of 0.78 to 0.81 over the baseline.

Tile Coding in Continuous State Space [Report] Fall 2021
Reinforcement Learning *Guide:* Prof. Abir Das
- Conducted an in-depth analysis of fixed tile coding's impact on learning optimization and rate by systematically varying the number of bins and tiles, comparing its performance to adaptive tile coding.

Automated Curation of a Large Scale Multimodal Dataset [Presentation] Fall 2021
Scalable Data Mining *Guide:* Prof. Sourangshu Bhattacharya
- Developed an automated solution to address the challenge of comprehending cumbersome E-manuals.
- Created a pipeline involving text-summarization, search query optimization and text-based video retrieval using PageRank to assign each section of the E-manual with a Youtube video to enhance comprehension.

Temporal Relational Reasoning in Videos [Presentation] Fall 2021
Deep Learning *Guide:* Prof. Abir Das
- Presented a ECCV 2018 paper explaining temporal dependencies using Temporal Relational Networks (TRN).
- Conducted additional experiments demonstrating the robustness of the TRN model against variations in video speed, thereby validating the model's adaptability and practicality in dynamic scenarios.

ApartWell: Apartment Management System [GitHub] Spring 2021
Database Management Systems *Guide:* Prof. Shamik Sural
- Developed a full stack website and Android app using SQLite database for apartment management
- Modelled the database using Entity-Relationship Model and designed the relational database schema.
- Included functionalities like visitor check-in, complaints booking, facilities booking etc.

OUTREACH/POSITIONS OF RESPONSIBILITY

Adobe Work Experience Program - Mentor Fall 2023
Mentored 10+ high school students in an ideathon promoting practical digital skills with an emphasis on GenAI.

Student Welfare Group, IIT Kharagpur - Mentor 2020 - 2021
Mentored a group of four freshers, offering guidance and support in both academic and extracurricular pursuits.

Rajendra Prasad (RP) Soccer Team, IIT Kharagpur - Captain, Member 2018 - 2022
Captained my dorm (RP) team in final year and consistently represented it since first year in Technology Gymkhana Championship. We were awarded gold in the second year and silver in the first year.

Encore: Technology Dramatics Society, IIT Kharagpur - Member 2018 - 2022
Performed three English plays, and produced and performed one Hindi street play.

Technology Filmmaking & Photography Society, IIT Kharagpur - Member 2018 - 2022
Scripted, directed, edited & acted in a short film called *Boysenberry* and successfully mentored the Freshers' Production

(c)ypher, School Computer Club of DPS Bhopal - President 2017 -2018
Led a team of 15+ students to organize (c)ync v4.0, one of the most prestigious tech symposium of Madhya Pradesh which were attended by 14 prominent schools of Bhopal. Conducted over 4 events testing coding, photography, gaming and electronic music production skills of students.

ACHIEVEMENTS

COMSNETS 2024 - Received the Best Poster Award among 48 posters at the 16th edition of the conference.

Prof. Mitra Cup 2022 - Recipient of the prestigious award for the best outgoing all-rounder from RP Hall.

ACM-ICPC 2021 - Ranked 2nd in IIT Kharagpur in Preliminary and 75th nationally in the Regional Round.

Facebook HackerCup 2020 - Global Rank 967 in Round 1 and qualified for Round 2.

IBM Green Hack - Awarded the first position at the IBM Green Hack 2020, a hackathon on climate change.

Competitive Examinations - Secured AIR 262 in JEE Advanced and 424 in JEE Main among 1.2 million students.

NSEC 2018 - Certificate of Merit by Indian Association of Physics Teachers for being in National Top 1%.

KVPY Fellowship - Awarded the prestigious fellowship by the Government of India with AIR 487.

NTSE Scholarship - Recipient of the coveted scholarship by the Government of India to top 2000 of the country.