Anirudh Atmakuru

(413) 210 – 4147 aa2886@cornell.edu LinkedIn GitHub Webpage

Education

Cornell University *August 2025 – Present*
Ph.D. in Computer Science — **Cumulative GPA: 4.15/4.0** *Ithaca, NY*
Coursework: Generative AI, High-performance and Parallel Computing

University of Massachusetts Amherst *August 2023 – May 2025*
MS in Computer Science — **Cumulative GPA: 3.90/4.0** *Amherst, MA*
Coursework: Advanced Machine Learning, Natural Language Processing (NLP/NLU), Reinforcement Learning, Intelligent Visual Computing, Methods of Applied Statistics, Systems for Data Science, Algorithms for Data Science

Birla Institute of Technology and Science, Pilani (BITS Pilani) *August 2018 – July 2023*
Dual Major - B.E. Computer Science and MSc. Mathematics — **Cumulative GPA: 9.15/10.0** *Pilani, India*
Coursework: Foundations of Data Science, Linear Algebra, Optimization, Statistics, Machine Learning (ML), Information Retrieval, Data Structures and Algorithms, Database Systems, Object Oriented Programming

Selected Publications (Google Scholar)

- **Does quantization affect models' performance on long-context tasks?** *EMNLP 2025*
  Anmol Mekala*, **Anirudh Atmakuru***, Yixiao Song, Marzena Karpinska, Mohit Iyyer
- **CS4: Measuring the Creativity of Large Language Models Automatically by Controlling the Number of Story-Writing Constraints** *EMNLP - WNU 2024*
  Anirudh Atmakuru*, Jatin Nainani*, Rohith Siddhartha Reddy Bheemreddy*, Anirudh Lakkaraju*, Zonghai Yao, Hamed Zamani, Haw-Shiuan Chang*
- **Deep Learning in Radiology for Lung Cancer Diagnostics: A Systematic Review of Classification, Segmentation, and Predictive Modeling Techniques** *Expert Systems with Applications 2024 (IF: 7.5)*
  Anirudh Atmakuru, Subrata Chakraborty, Oliver Faust, Massimo Salvi, Prabal Datta Barua, Filippo Molinari, UR Acharya
- **Do Deep Learning Models Mimic Human Personality Traits? – An Empirical Study** *ACIS 2022*
  Anirudh Atmakuru, Subrata Chakraborty
- **Hidden and Face-like Object Detection using Deep Learning Techniques – An Empirical Study** *DICTA 2022*
  Anirudh A, Subrata Chakraborty
- **Prediction of Video Game Development Problems Based on Postmortems using Different Word Embedding Techniques** *ICON 2021*
  Anirudh A, Aman Raj Singh, Anjali Goyal, Lov Kumar, N L Bhanu Murthy
- **Mathematical Modeling and the Transmission Dynamics in Predicting COVID-19 – What Next in Combating the Pandemic?** *Infectious Disease Modelling 2020*
  A Anirudh

Projects

Hallucination Memorization in Outcome-based Post-Training March 2026 – August 2026

- Identified an unstudied failure mode of **RLVR**: outcome-based rewards drive LLMs to memorize self-generated hallucinations from positively rewarded rollouts, and proposed a **memorization score** isolating this from RL-induced distribution sharpening. [Under review]
- Demonstrated that the effect holds across models, architectures, and levels of prior knowledge, is strongest on hard training questions, persists over training, survives a KL penalty, and that models surface memorized hallucinations even when allowed to abstain.

Characterizing Epistemic Uncertainty in LLMs for Knowledge-Intensive Tasks September 2025 – February 2026

- Comprehensively studied verbalized epistemic uncertainty in **reasoning** LLMs for **long-form knowledge-intensive QA**, introducing a fact-level taxonomy of reasoning behaviors and uncertainty expressions, shifting analysis from answer-level to claim-level confidence. [Under review]
- Showed that verbalized uncertainty correlates with factuality and fact inclusion in responses despite RLVR not incentivizing this, and proposed a training-free operationalization strategy exploiting both signals to improve factuality by over 8%.

Impact of LLM Quantization on Text Quality August 2024 – May 2025

- Conducted the first systematic evaluation of quantized LLMs for long-context reasoning and long-form text generation, benchmarking 30 model-quantization pairs, noting a drop of over **10%** in quality for the first time in literature. [Published] [**GPTQ, AWQ, Perplexity, BLEU, Llama, Qwen, GPU, CUDA**]
- Found that **8-bit** quantization preserves model performance with minimal degradation, while **4-bit** quantization leads to severe drops (up to **59%**) on multilingual and long-context tasks.

Prompt Specificity - Unlocking Secrets of LLM Behavior January 2024 – August 2024

- Designed and implemented an evaluation benchmark to assess the trade-off between instruction satisfaction and long-form text generation quality, noting over **30%** depreciation in LLM performance due to the trade-off. [Published]
- Analyzed the impact of prompt specificity on 7B models to understand training data properties and their effects on different stages of model training, including **SFT** and Reinforcement Learning with Human Feedback (**RLHF**).

Undergraduate Thesis - Computer Vision August 2022 – December 2022

- Developed **CNN**-based classifiers and detectors to differentiate faces from face-like objects and identify face-like objects hidden amidst background clutter with an accuracy of over **80%** in different settings. [Published]
- Conducted psychological image tests, including left-brain/right-brain and Rorschach inkblot tests, to identify biases in deep learning and transfer learning models and evaluate their mimicry of human personality traits. [Published]
- Created a transfer learning framework to classify non-verbal vocalizations from minimally speaking autistic individuals by considering audio data as images using Mel spectrograms, achieving a SOTA F1 score of **0.79**, compared to the previous best of 0.52. [**Multimodal, Deep Learning, DenseNet, ResNet, VGG-19, GoogleNet, MobileVNet**] [Published]

BioNLP - Self-rewarding LLMs for Medical Question Answering May 2024 – August 2024

- Implemented the self-rewarding pipeline on Biomistral-7B, tailored for medical QA systems, to enhance AI responsiveness and accuracy in healthcare applications, outperforming GPT-4 by a **20.44%** win rate. [**Synthetic data, LLM-as-Judge, RLHF, RLAIF, PPO, DPO**]

Dynamic Fine-Tuning of LLMs Using RLHF for Improved Alignment February 2024 – May 2024

- Presented a proof-of-concept to enhance LLM alignment with human values through dynamic prompt generation based on intermediate model performance, resulting in improved model accuracy, optimization, and alignment.
- Introduced and evaluated two novel training approaches - Always-On-Policy and Starts-On-Policy, with generations being preferred **61.85%** of the time compared to traditional RLHF training.

Data Science Internship at Epsilon, India January 2022 – April 2022

- Built a **transformer** model to generate diverse email subject lines based on a theme or starting word(s), increasing advertisement email open rates by over **11%.**
- Generated 10 semantically similar subject lines by considering it an English-English neural machine translation task.
- Built an **LSTM** model to predict open rates of emails that use the generated subject lines with a **91% accuracy**.

Emotionally Aware Conversational Agent for Mental Health August 2021 – December 2021

- Built a multi-faceted, emotionally aware chatbot combining rule-based, retrieval-based (**LSTMs, GRUs**), and generative methodologies (**seq2seq encoder-decoder**-based text generator).

Prediction of Video Game Development Problems using Word Embeddings July 2021 – October 2021

- Built machine learning, ensemble learning, and multi-layer perceptron models, achieving an AUC score of **0.97**, to classify video game development problems into different groups based on the quote and problem description. [Published]
- Performed a comparative analysis on five word embedding techniques, SMOTE, three feature selection techniques, five ML classifiers, and three ensemble learners. [Published] [**PCA, LDA, ANOVA, KNN, Naive Bayes, SVM, SMOTE**]

Fellowships

- **Foundational AI PhD fellow** for the 2025-2026 academic year.

Responsibilities

- **Reviewer** for the International Journal of Imaging Systems and Technology (2024, 2025), Second Workshop on Customizable NLP (ACL 2026).
- **Teaching Assistant** for the courses Elementary Real Analysis (Fall 2020), Probability and Statistics (Spring 2020).
- **Grader** for the COMPSCI 696DS: Industry Mentorship Independent Study - Data Science course (Spring 2025).

Technical Skills (GitHub)

Languages: Python, C, C++, Java, MATLAB, JavaScript, R.
Python Libraries: Pandas, Numpy, Scipy, Matplotlib, TensorFlow, PyTorch, Scikit-learn (SKLearn), HuggingFace, vLLM, verl.
Developer Tools and Frameworks: Azure, Terraform, HTML, CSS, SQL, SPSS Statistics, Amazon Web Services (AWS), MongoDB, MySQL, Cassandra, PostgreSQL, React, Angular, Git, Docker, Linux, Agile, Kubernetes, Jupyter.