ANSH ARORA

📞 +1-(413)-466-4786   ✉ ansharora@umass.edu   in ansharora   ○ ansharora7

EDUCATION

University of Massachusetts Amherst   **Amherst, Massachusetts, U.S.A**
M.S in Computer Science, GPA - 3.9/4.0   Expected Graduation: May, 2026

Indian Institute Of Information Technology, Guwahati   **Guwahati, Assam, India**
B.Tech in Computer Science and Engineering, CGPA - 9.15/10.0 (Ranked 5/138)   May, 2024

RELEVANT EXPERIENCE

Google Deepmind   **February 2025 – Present**
Graduate Researcher | Python, PyTorch, HuggingFace, Jupyter, Deep Learning
- Researching novel meta-optimization strategies for deep learning to enhance model generalization and efficiency.
- Developing an adaptive model ensembling approach to improve training efficiency.

MAQ Software   **December 2023 – May 2024**
Associate Software Engineer | Tensorflow, Pandas, MySQL, MS Azure, vLLM, Data Engineering and Pipelining
- Enhanced LLM-based chatbot functionality by enabling PDF/CSV file uploads and text-based query retrieval, leveraging vLLM for model hosting and NLP techniques for document parsing to improve user accessibility.
- Built propensity prediction models for Microsoft's FastTrack Core ML Team, achieving AUROC 0.8 on a dataset with 20M+ rows and 100+ columns to optimize tenant conversion.

University College London   **August 2023 – February 2024**
Research Assistant | Pytorch, Natural Language Processing (NLP), LLMs, AI security
- Built an end-to-end inference-time defense pipeline against backdoor attacks in NLP models using model merging, reducing attack success rates by 75% outperforming all state-of-the-art defenses; work accepted in *ACL 2024*.
- Evaluated defense effectiveness across architectures through extensive experiments on SST-2, QNLI, Amazon, and IMDB datasets using BERT and RoBERTa models, as well as LLMs like Llama2-7B and Mistral-7B.

PUBLICATIONS & AWARDS

- ACL 2024 - Arora, Ansh, et al. "Here's a Free Lunch: Sanitizing Backdoored Models with Model Merge." (2024)

- Recipient of DAAD-WISE Scholarship (2023), awarded to ∼150 students for research internships in Germany.

PROJECTS

Large Language Model as Teacher for Extreme Classification (LMTX)
Aalto University, Finland | C++, Extreme Classification, Zero-shot setting, NLP, LLMs
- Designed an Extreme Multi-label Classification system for zero-shot tagging, improving document-label correlation analysis using advanced feature encoding.
- Enhanced document-label correlation discovery by leveraging LLMs as teachers to guide the feature encoder, leading to more accurate zero-shot predictions.

Multi-label Generalized Zero-shot Learning for Diagnosis of Chest Radiographs (GZSL-X-ray)
IIT Jodhpur | OpenCV, Keras, Computer Vision (CV), Medical Imaging, Similarity Search
- Identified abnormalities in Chest X-ray images using variational autoencoders in a Multi-label Generalized Zero-shot setting, achieving AUROC 0.73 and F1-score 0.65 on the NIH Chest X-ray dataset.
- Built an end-to-end pipeline with EfficientNet-b4 for visual extraction and BioBERT for textual embeddings, using similarity measures for embedding matching.

TECHNICAL SKILLS

Languages: C, C++, Python, Java, R	**Version Control & Dev Tools:** Git, Vim
GPU & Parallel Computing: CUDA	**Data Pre-Processing:** Pandas, NumPy, Scikit-learn
Databases: MongoDB, MySQL, PostgreSQL	**Cloud Services:** AWS, MS Azure
Deep Learning Frameworks: TensorFlow, PyTorch,	**Productivity Tools:** MS Office
Hugging Face, Keras, OpenCV	