ANSH ARORA

+1-(413)-466-4786 ansharora@umass.edu linkedin.com/in/anshkarora github.com/ansharora7

EDUCATION

University of Massachusetts Amherst — **Amherst, Massachusetts, U.S.A**
M.S in Computer Science, GPA - 3.96/4.0 — 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

Thales (Digital Identity & Security Division) — **May 2025 – December 2025**
Research Engineer Intern | Biometrics, Computer Vision, Large-Scale Training, ONNX Deployment

- Designed and led development of a CNN-based dual-branch fingerprint identification architecture with cross-attention for inter-branch feature fusion.
- Trained large-scale models on 1.2M+ subjects (12M images) using Fully Sharded Data Parallel (FSDP)/Distributed Data Parallel (DDP) across 5+ GPUs, with gradient checkpointing to improve memory efficiency and reduce training cost.
- Built augmentation pipelines incorporating region-of-interest (ROI) masking, sensor noise simulation, and synthetic data generation to improve model robustness.
- Generated 100K+ synthetic slap and rolled fingerprints using Generative Adversarial Networks (GANs) and diffusion models, improving generalization across sensors.
- Implemented scalable evaluation pipelines and benchmarked models on a 1M+ gallery using Facebook AI Similarity Search (FAISS), performing >100B pairwise comparisons.
- Optimized models for CPU-efficient inference using ONNX and OpenVINO, integrated into production-style workflows.
- Outperformed internal company baseline by +7% rank-1 accuracy and -6% false rejection rate (FRR) on proprietary benchmarks.
- Authored 5 patents on neural architectures, augmentation strategies, and synthetic-to-real data transferability.

Google DeepMind — **February 2025 – May 2025**
Graduate Student Researcher | Python, PyTorch, HuggingFace, Jupyter, Deep Learning, Vision Transformers

- Designed a data-free, gradient-based model souping algorithm that ensembles pretrained model weights without requiring any access to training data.
- Eliminated the need for retraining or fine-tuning by directly optimizing in weight space, making the approach broadly applicable across tasks.
- Reduced runtime by 90.5% and memory usage by 63.8% compared to standard model souping methods, enabling much faster experimentation.
- Validated the method across ImageNet, CIFAR, and GLUE benchmarks, including successful demonstrations of mixing domain expert models.

MAQ Software — **December 2023 – May 2024**
Associate Software Engineer | TensorFlow, Pandas, MySQL, MS Azure, vLLM, Data Engineering and Pipelining

- Enhanced Large Language Model (LLM)-based chatbot functionality by adding support for PDF and CSV file uploads with text-based query retrieval.
- Leveraged vLLM for scalable model hosting and applied natural language processing (NLP) techniques for efficient document parsing.
- Developed propensity prediction models for Microsoft's FastTrack Core ML Team, training on a dataset of 20M+ rows and 100+ columns.
- Achieved an AUROC of 0.8, enabling optimization of tenant conversion strategies.

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 techniques.
- Introduced a data-free approach that neutralizes malicious triggers at inference without requiring retraining or access to original training data.
- Reduced attack success rates by 75%, outperforming all existing state-of-the-art defenses, with results accepted at *ACL 2024*.
- Evaluated effectiveness extensively on SST-2, QNLI, Amazon, and IMDB datasets using both BERT and RoBERTa architectures.
- Validated scalability to large language models (LLMs) such as Llama2-7B and Mistral-7B, establishing the method as a widely recognized baseline in backdoor defense research.

RESEARCH & RECOGNITIONS

- Member of Organizing Committee – Anti-BAD: Anti-Backdoor Challenge for Post-Trained Large Language Models, accepted as a competition at IEEE SaTML 2026.
- ICLR 2026 – Li, Weijun, Arora, Ansh, et al. "Defending Deep Neural Networks against Backdoor Attacks via Module Switching."
- ACL 2024 – Arora, Ansh, et al. "Here's a Free Lunch: Sanitizing Backdoored Models with Model Merge."
- 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
- **Deep Learning Frameworks:** TensorFlow, PyTorch, Hugging Face, Keras, OpenCV
- **GPU & Distributed Computing:** CUDA, FSDP, DDP, Gradient Checkpointing
- **Inference & Deployment:** OpenVINO, ONNX
- **Data Pre-Processing:** Pandas, NumPy, Scikit-learn
- **Databases:** MongoDB, MySQL, PostgreSQL
- **Cloud Services:** AWS, Microsoft Azure
- **Version Control & Dev Tools:** Git, Vim
- **Productivity Tools:** MS Office