Anirudh Som

Contact Info: **Email:** contact.anirudh.som@gmail.com **Work Authorization:** U.S. Permanent Resident **LinkedIn** **Google Scholar**

Summary: Experienced AI/ML lead with over 10 years of expertise in machine learning, deep learning, and NLP systems. Currently serving as the technical lead and senior ML scientist at SRI International, specializing in building NLP pipelines for real-world business and government applications. Proven ability to design scalable solutions, mentor junior scientists, and translate complex algorithms into actionable products.

Skills:

Programming – Python, Matlab, XML, Latex

Libraries/Frameworks – PyTorch, TensorFlow, Keras, Scikit-learn, Pandas, NumPy, OpenCV

NLP & Deep Learning – Transformers, Large language models, POS tagging, Tokenization, OCR

DevOps & Deployment – Docker, Kubernetes, Gitlab, Artifactory, AWS

Data Handling – Document extraction, Financial datasets, Unstructured to structured data conversion

Education:

Ph.D., Electrical Engineering – *Arizona State University*, USA **2020**

Experience:

[1] **Advanced Computer Scientist** – *SRI*, Princeton, NJ **2021 – Present**

- *NLP Systems*: Designed and led the development of transformer-based NLP models for cross-cultural dialogue understanding and human-machine interfaces under DARPA and NSF programs.
- *Document Intelligence*: Developed LLM-based systems for extracting structured insights from unstructured medical and educational documents. Built OCR + NLP pipelines for scanned inputs.
- *Mentorship & Leadership*: Acted as technical lead for multiple federal and commercial AI projects. Supervised 2–4 junior ML engineers and interns across project lifecycles.
- *Commercial Use Cases*: Delivered NLP prototypes and AI solutions to commercial clients, including L'Oréal and Honda.
- *Product Collaboration*: Integrated ML modules into platforms, ensuring performance monitoring and maintenance.

[2] **Previous Internship Experience** **2015 – 2020**

- *SRI*: Developed deep learning models for multimodal behavior analysis and automated feedback generation from unstructured classroom data.
- *Roche Diagnostics*: Built deep learning pipelines for image segmentation and integrated AI models into platforms for processing medical documents. Developed a federated learning framework for privacy-preserving medical image analysis, addressing data security in distributed systems.
- *Lawrence Livermore National Laboratory*: Designed time-series deep learning algorithms for health signal modeling, addressing data imbalance and automation of diagnostics.
- *Mayo Clinic*: Created ML algorithms to identify cancerous regions in CT/MRI images, demonstrating applied pattern recognition in medical datasets.

Patents:

[1] Machine Learning Model Prompt Demonstration Selection, 2025.
[2] Federated Learning System for Training Machine Learning Algorithms Without Accessing Patient Data to Protect Patient Privacy, 2022.
[3] Automated Collaboration Skills Assessment, 2021.

Recent Publications:

[1] **A. Som**, K. Sikka, H. Gent, A. Divakaran, A. Kathol, D. Vergyri
Demonstrations Are All You Need: Advancing Offensive Content Paraphrasing using In-Context Learning
Accepted at the Association for Computational Linguistics (ACL) Findings, 2024.

[2] **A. Som**, S. Kim, B. Lopez-Prado, S. Dhamija, N. Alozie, A. Tamrakar
Automated Student Group Collaboration Assessment and Recommendation System Using Individual Role and Behavioral Cues
Accepted at the Frontiers in Computer Science Journal, 2021.

Award & Recognition:

- "Research Recognition Award" – CVT Group at SRI International **2022 – 2024**
- "Facilitators' Choice Award" - NSF STEM for All Video Showcase **2021**
- "Most Innovative Award" - Roche **2019**